

SEARCH REQUEST FORM

2-631

Requestor's Name: _____ Serial Number: _____

Date: _____ Phone: _____ Art Unit: 9E11

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

STAFF USE ONLY

Date completed: 2/1/97

Searcher: Sheppard

Terminal time: _____

Elapsed time: _____

CPU time: _____

Total time: _____

Number of Searches: _____

Number of Databases: _____

Search Site

_____ STIC

_____ CM-1

_____ Pre-S

Type of Search

_____ N.A. Sequence

_____ A.A. Sequence

_____ Structure

_____ Bibliographic

Vendors

_____ IG

_____ STN

_____ Dialog

_____ APS

_____ Geninfo

_____ SDC

_____ DARC/Questel

_____ Other

W P E R L H (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Thu Feb 18 14:40:49 1999; MasPar time 2624.06 Seconds
1532.792 Million cell updates/sec
Abular output not generated.

Title: >US-08-934-254-26
Description: (1-1702) from US08934254.seq
Perfect Score: 1702
N.A. Sequence: 1 CCCCAAAATTTTCATTGTT.....TTTTGGTAAAAAAAAA 1702
Comp: GGGGTTTTTAAAGTAACAA.....AAAAACCTTTTTTTTTTT

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 602357 seqs, 1181590623 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: emb157
1:em_ba 2:em_fun 3:em_htg 4:em_hum1 5:em_hum2 6:em_in
7:em_om 8:em_or 9:em_ov 10:em_pat 11:em_ph 12:em_pl
13:em_ro 14:em_un 15:em_vi
genbank110
16:gb_bal 17:gb_ba2 18:gb_htg 19:gb_in 20:gb_om 21:gb_ov
22:gb_pat 23:gb_ph 24:gb_pl1 25:gb_pl2 26:gb_pr1
27:gb_pr2 28:gb_pr3 29:gb_ro 30:gb_st 31:gb_sts 32:gb_sy
33:gb_un 34:gb_vi

Statistics: Mean 11.427; Variance 5.921; scale 1.930

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Match	Length	ID	Description	Pred. No.
1	399	23.4	110149	25	ATAC005397 Arabidopsis thaliana c	4.77e-288
2	393	23.1	1705	25	ATJ4161 Arabidopsis thaliana m	3.71e-283
3	385	22.6	1610	25	BNJ4160 Brassica napus mRNA fo	1.22e-276
4	340	20.0	1591	25	HACTB5RN H.annuus mRNA for extr	4.63e-240
5	172	10.1	1685	22	AR020904 Sequence 4 from patent	1.78e-105
6	172	10.1	1685	22	Sequence 4 from patent	1.78e-105
7	172	10.1	1687	25	Borago officinalis del	1.78e-105
8	91	5.3	7218	22	Sequence 14 from paten	2.20e-43
9	46	2.7	965	22	Sequence 22 from paten	7.60e-12
10	43	2.5	281	24	RYENRA Secale cereale nitrate	6.13e-10
11	42	2.5	10772	19	AF012089 Drosophila melanogaste	2.59e-09
12	40	2.4	965	22	Sequence 22 from paten	4.48e-08
13	40	2.4	10772	19	AF012089 Drosophila melanogaste	4.48e-08

C	14	37	2.2	7218	22	I66494	Sequence 14 from paten	2.92e-06
C	15	38	2.2	74371	27	AC005369	Homo sapiens chromosom	7.36e-07
C	16	37	2.2	74371	27	AC005369	Homo sapiens chromosom	2.92e-06
C	17	36	2.1	215	22	I28278	Sequence 5 from patent	1.14e-05
C	18	35	2.1	285	24	BLYNRB	Hordeum chilense nitra	4.41e-05
C	19	35	2.1	285	24	BLYNRC	Hordeum lechleri nitra	4.41e-05
C	20	35	2.1	285	24	BLYNRD	Hordeum pusillum nitra	4.41e-05
C	21	36	2.1	1801	24	MZENADNR	Zea mays nitrate reduc	1.14e-05
C	22	35	2.1	3076	24	CUCNITRA	C.maxima nitrate nat7 g	4.41e-05
C	23	35	2.1	3538	24	HVNAR7	Hordeum vulgare nat7 g	4.41e-05
C	24	36	2.1	4985	24	ZMU20450	Zea mays nitrate reduc	1.14e-05
C	25	35	2.1	7602	24	VCNITA	V.carteri gene for nit	4.41e-05
C	26	34	2.0	215	22	I28278	Sequence 5 from patent	1.67e-04
C	27	34	2.0	2000	24	OSNIA134	Rice nial gene for nit	1.67e-04
C	28	33	1.9	282	24	BLYNRA	Hordeum bogdanii nitra	6.22e-04
C	29	33	1.9	285	24	BLYNRF	Hordeum stenostachys n	6.22e-04
C	30	32	1.9	2126	24	MZENAR	Maize NADH:nitrate red	2.27e-03
C	31	33	1.9	5499	24	LJNIA	L.japonicus NIA gene.	6.22e-04
C	32	31	1.8	273	24	ESTNRA	Avena strigosa nitrate	8.15e-03
C	33	31	1.8	540	22	E13060	cDNA encoding cytochro	8.15e-03
C	34	31	1.8	565	22	E04076	gDNA encoding envelope	8.15e-03
C	35	30	1.8	3811	26	HUMIMUCA	Human mucin 2 (MUC2) m	2.86e-02
C	36	30	1.8	15720	27	HUMMUC2X	Homo sapiens intestina	2.86e-02
C	37	30	1.8	112309	27	AC003025	Human Chromosome 11p12	2.86e-02
C	38	30	1.8	170743	18	AC004228	*** SEQUENCING IN PROG	2.86e-02
C	39	30	1.8	177014	27	AC004063	Homo sapiens chromosom	2.86e-02
C	40	29	1.7	954	24	CVNIREDBH	C.vulgaris mRNA for ni	9.84e-02
C	41	29	1.7	38962	16	SC2F1	Streptomyces coelicolo	9.84e-02
C	42	29	1.7	75342	19	CEY47H9C	Caenorhabditis elegans	9.84e-02
C	43	29	1.7	110000	18	CEY53C10	Caenorhabditis elegans	9.84e-02
C	44	29	1.7	181098	27	AC004216	Homo sapiens 12q24 PAC	9.84e-02
C	45	29	1.7	231298	18	AC005921	*** SEQUENCING IN PROG	9.84e-02

ALIGNMENTS

RESULT	1	ATAC005397	110149 bp	DNA	PLN	06-OCT-1998
LOCUS		Arabidopsis thaliana chromosome II	BAC T3F17	genomic sequence,		
DEFINITION		complete sequence.				
ACCESSION		AC005397				
NID		93702315				
KEYWORDS		HTG.				
SOURCE		thale cress.				
ORGANISM		Arabidopsis thaliana				
REFERENCE		1 (bases 1 to 110149)				
AUTHORS		Rounsley, S.D., Lin, X., Kaul, S., Shea, T.P., Fujii, C.Y., Mason, T.M., Shen, M., Ronning, C.M., Fraser, C.M., Somerville, C.R. and Venter, J.C.				
TITLE		Arabidopsis thaliana chromosome II BAC T3F17 genomic sequence				
JOURNAL		Unpublished				
REFERENCE		2 (bases 1 to 110149)				
AUTHORS		Rounsley, S.D. and Lin, X.				
TITLE		Direct Submission				
JOURNAL		Submitted (07-AUG-1998)				
REFERENCE		3 (bases 1 to 110149)				
AUTHORS		Rounsley, S.D.				
TITLE		Direct Submission				
JOURNAL		Submitted (06-OCT-1998)				
COMMENT		Medical Center Dr., Rockville, MD 20850, USA				
		On Oct 6, 1998 this sequence version replaced gi:3603083.				
		Address all correspondence to:				
		Steve Rounsley				
		The Institute for Genomic Research				
		9712 Medical Center Dr.,				
		Rockville, MD 20850,				
		USA				
		e-mail: rounsley@tigr.org				
		BAC clone T3F17 is from Arabidopsis chromosome II and is contained				

/note="hypothetical protein"
/codon_start=1
/db_xref="PID:g3702319"
/translation="MADYQNPVLRKPYRDPNMSPPPPPPPIQQQPMKXAVPMT
SYRKKRRSCRCFCGCCICITLVLFILLLVAVFYLFDFPKLPTFSLASRLDGF
KLADPDGASLSATAVARVEMKPNKSLVYVYGNVAVDLVSVGNDGTGEMTNGF
ROGPKNSTSVKVTETVKNLVERGLAKRLAQQFQSKDLVNVVAKTKVGLGVGGIKIG
MLAVNRCGGVSLNKLDTPSPKILNTLWKYLIISN"
mRNA join(.15933..16138,16274..16489,16586..16719,16820..16939,
17042..17131,17227..17313,17401..17480,17583..17768)
/gene="T3F17.6"
15933..17768
/gene="T3F17.6"
join(.15933..16138,16274..16489,16586..16719,16820..16939,
17042..17131,17227..17313,17401..17480,17583..17768)
/gene="T3F17.6"
/codon_start=1
/product="putative TGF-beta receptor interacting protein"
Note: remainder of annotations omitted.

Query Match 23.4%; Score 399; DB 25; Length 110149;
Best Local Similarity 67.3%; Pred. No. 4.77e-288;
Matches 835; Conservative 0; Mismatches 400; Indels 6; Gaps 5;

Db 44741 GCTTGAATCGCGGTTCTTCAAGTCTTAATCGTCCACATFAGC-T--TCCACCCAC 44797
Cp 1336 GCCTGAACCGCGCATCCCTCAGCGTCCGAATTGCTCTGACATTAGCGTCGTCCCAAAAC 1277
Db 44798 GAAAGACTCTATACGTAGATTATGTTCTTACAAAGTCTTAAACACCGCGGATCT 44857
Cp 1276 CCGAAGTCTCTATACCGCATCCCGTGTCTTCTTACAAAGTCTTAAACACCGCGGATCT 1217
Db 44858 GTCCGGAGTACAAACGAGGAGCGGAGGAAACAAATGATGCTCTAACTGAACTGCAAT 44917
Cp 1216 TTCTAGTCTGCCACCGCGGAGCTTAGGAACAAGTGGTCTCCAACCTGGAACCTGCAGC 1157
Db 44918 CCACCAAGAACCAATCCATAAAGTCTACACGATATATCAAGCGTACCGCGGTTGT 44977
Cp 1156 CCACCAAGAACCAATCCATCCAGGTGGCGACGTGATATCGATCGTCCCTTTCGTCTGC 1097
Db 44978 TTCTCAACCAATGTTTCCGTTTGGTGACCGGTATACACATCTGTGCGAAATGTTT 45037
Cp 1096 TTCTCGAACCAATGTTTCCGTTTGGTGACCGGTATACACATCTGTGCGAAATGTTT 1037
Db 45038 AACAGACTCAACATCTTGAATCCCGTGCAGCGGATCTACGAGACGAGCAATGAAT 45097
Cp 1036 AGCGTGAACCTGGAGTCTGGATCGCGTGCACCGCAAGCTGATGAGGACGAGCAACCGAAC 977
Db 45098 CTCCTTTGCCAATTCGGAAGCAAGCAGATTAAGAGGAGCAACCATGTCCAGAAACTAAA 45157
Cp 976 CGTTACGCGCAGTTCGGGACACAGATACGAGAGCGGAAACACGATCCAGAAACCGCG 917
Db 45158 ATTCGCGCATGTTTCAAGGCGCGATCTGGGACGTGGCGTTTTCGAGATAGCAATAGGAAT 45217
Cp 916 ATACCATTAAGTTTAGAGCGCGTTCAGGACGTGCGCTGTGTGAGGACCAATAAAAG 857
Db 45218 GTTTGGATAGAGGTTGATCTTCCACGCAATTAATCTGGGTAAATGTCACATGTTGG 45277
Cp 856 GTCTGGATAGAGGTTGATCTTCCACGCAATTAATCTGGGTAAATGTCACATGTTGG 797
Db 45278 TAGCTGATTAGAGTTCAGGTAGAGGATCAATTAATTCCTGCCATAGAAACGTGAC 45337
Cp 796 TAGCTGATTAGAGTTCAGGTAGAGGATCAATTAATTCCTGCCATAGAAACGTGAC 737
Db 45338 GTCATCAATTTAGAAATTTGTAGAGACGCGGAAGATCGGGATGTGTGTAGATCCGGA 45397
Cp 736 GTGATGAGTTAGAGTTCGGGTGACACGCGCAATACGGGATGTCTGGAGTCCGGG 677
Db 45398 TCGTGTCAAGGCTGTTTACAGCGATGTTGTGACGCTGTGTCGTCATTTCCACACCGG 45457
Cp 676 TCGTGTCAAGGCTGTTTACAGCGATGTTGTGACGCTGTGTCGTCATTTCCACACCGG 617

Db 45458 ATTGAGATCCCGGTGAGACAGTTACCGGAGAGAGCTGGATTAGTTTGTACACGGTTG 45517
Cp 616 ATGTGATCCCGGTAGGATGTGCTCTATGATGATTCGTCATTTCTGTATCCACGG 557
Db 45518 GTTGACGTACCGGTAGTACCGGAATCGTGGCGAGCTAAGCGCTTTGGATCCAGAGG 45577
Cp 556 GTTGACATCACTGTAATGCGCGGAGTATGCGCCACATAGCGGCTTGGATCCACAGC 497
Db 45578 AGACCGGATAAAGCGGAGATGAGGTGAGCCACAGATGCTGGTACAGCCAAAACGCGG 45637
Cp 496 AAGCCAGCAGTGGCGCCGACAGCATGTGAATCCGAGGACTCCGACGCCACGACGCG 437
Db 45638 TAGAGAACCGCGGAGCATGACCGGACGACGTGATGTGTAAGAGTCACTGACGCT 45697
Cp 436 TAGAGATTTGCGGCATCATGACCGCAACGCCGACGACGATGATGTTGGTGGGCC 377
Db 45698 TTTTGTGAGAGGCGCGTTTGGAGAACTCGCGGCTAAACGACGCTAGTCACGTTGAG 45757
Cp 376 TTTCTCGAAGATCCCGGACCGGACATCTCGTTCAAAGCCTCCGTTAGTCTTTGAG 317
Db 45758 ACCTCTGACAGCTGTGCTCTCTACGT-G-A-TAGCATTGTGAAGCTTTTCTAAGTGG 45814
Cp 316 ATCTCCGACATCTCGAGTCTCTGAGGTAGTAGTAGCCGTTGAGAGCGGATCCAGATGC 257
Db 45815 TGCCATCGGTTCCCGGATGTAAGCGATGAACCGTCCGTTGACGCTTTGGCCGCGGAGA 45874
Cp 256 CGCCACGCGTCCCGGATGTAAGCGATGAACCGTCCGTTGACGCTTTGGCCGCGGAGA 197
Db 45875 TTGAGATCGTCTGCTCCGCTCCCGGATGAGATTAAACCAATCGGAAACGCTGCTAACT 45934
Cp 196 CTGAGGACGCGGACCTCCGCGCGGCTGCTCCGCGCGCCACCGAGAGCAGTCTGTAGAC 137
Db 45935 TTACCTTGAATCGAAATCCATAAATCTCCAGTGTGTTGTG 45975
Cp 136 TTGCTGTGATGAGATCCAGATGCGCGGACTGTTGTG 96

RESULT 2
LOCUS ATAJ4161 1705 bp mRNA PLN 31-OCT-1998
DEFINITION Arabidopsis thaliana mRNA for delta-8 sphingolipid desaturase.
ACCESSION AJ224161
NID g3819709
KEYWORDS delta-8 sphingolipid desaturase; fusion protein; sld1 gene.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta;
Magnoliophyta; eudicotyledons; Rosidae; Capparales; Brassicaceae;
Arabidopsis.
REFERENCE 1 (bases 1 to 1705)
Sperling,P., Zahringer,U. and Heinz,E.
A sphingolipid desaturase from higher plants. Identification of a
new cytochrome b5 fusion protein
J. Biol. Chem. 273 (44), 28590-28596 (1998)
99003197
REFERENCE 2 (bases 1 to 1705)
Sperling,P., Zahringer,U. and Heinz,E.
Identification of cDNAs coding for a sphingolipid desaturase from
higher plants
Unpublished
REFERENCE 3 (bases 1 to 1705)
Sperling,P.
Direct Submission
Submitted (13-FEB-1998) Sperling P., Physiologie, Institut fuer
Allgemeine Botanik der Universitaet Hamburg, Ohnhorststrasse 18,
D-22609 Hamburg, FRG
Related sequence: X87143.
COMMENT Location/Qualifiers
FEATURES
source 1..1705
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/clone_lib="lambda ZAP"

Qy	723	TTCAACTCCATCACTCGGTCCTTCTATGGCCGAGTCTCGAAATTCGACGAAGTGGCACGG	782
Db	907	TTCTTAGTCAGCTATCAACACTTTTACTTATTATCCAGTTATGCTTTGGAAGAATCAAT	966
Qy	783	TTCTTAGTCAGCTACCAAGCACTGACCTACTACCGGTGATGATCTTCGGCCGAGTCAAC	842
Db	967	CTCTTCAATCAAAAGTTTCTCTTCTGCTCTTCTCCAAAGGTGAAGTACAGATCGTGTCTTA	1026
Qy	843	CTCTTCAATCCAGACCTTTTATTGCTCTCAACAGGCGACGTCCTGACCGCGCTCTA	902
Db	1027	AACTTCGCGCGAATCTAGTCTCTCTGGACTTGGTTCCTCACTCTTAGTCTCATGCTTACCA	1086
Qy	903	AACTTAATGAGTATCGCGGTTTCTGACGTGGTTCCTGCTCTTCGTATCTTGTCTCCCG	962
Db	1087	AACTGGGCTGAGAGATCTTCTTCGTCTTCAAGCTTCACCGTCACGGCGCTTCAACAC	1146
Qy	963	AACTGGGCTGAACGGTTCGGGTTCGTCTCATAGCTTTGGGTGACGGGATCCAGCAC	1022
Db	1147	ATTCAATTCAGCTTTAACCAATTTGCTGTGATGTGTACGTTGGTCCACCCACCGGTAGC	1206
Qy	1023	GTCCAGTTTCAGCTTCAACCACTTCTCCGGGACACATACGTGGGCCCCCAAGGGCGAC	1082
Db	1207	GACTGGTTCGAGAGCAAGCGCGGGAACAATCGATATCTTGTAGATCATATCATGGAT	1266
Qy	1083	AACTGGTTCGAGAGACAGACAAGGAGCAGCATCATACGTGCCACCGTGGATGGAC	1142
Db	1267	TGGTCTTTGGTGATATACAGTTTCAGCTTGAGCATCATTTGTTTCCCTCGCTTACCTCGT	1326
Qy	1143	TGGTCTTTGGTGAGTTCAGTTTCCAGTTTGAGACCACTTGTTCCTTAGCTGCGCGT	1202
Db	1327	TGCCATCTCCGGAAGTTTCTCCGGTGGTTCAAGAGGTTTCAAGAAGCATATCTTCCG	1386
Qy	1203	GGCAGCTTAGAGAGATTGCGCCCTTGGCTCGGACCTTGTGAAGAAGCACGGGATGCCG	1262
Db	1387	TATAGGAG 1394	
Qy	1263	TATAGGAG 1270	
RESULT 3 BNAJ4160 1610 bp mRNA PLN 31-OCT-1998			
DEFINITION Brassica napus mRNA for delta-8 sphingolipid desaturase.			
ACCESSION AJ224160			
NID g3819707			
KEYWORDS delta-8 sphingolipid desaturase; fusion protein; sid1 gene.			
SOURCE rapeseed			
ORGANISM Brassica napus			
REFERENCE 1 (bases 1 to 1610)			
AUTHORS Sperling, P., Zahringer, U. and Heinz, E.			
TITLE A sphingolipid desaturase from higher plants. Identification of a			
new cytochrome b5 fusion protein			
J. Biol. Chem. 273 (44), 28590-28596 (1998)			
JOURNAL 99003197			
MEDLINE 99003197			
REFERENCE 2 (bases 1 to 1610)			
AUTHORS Sperling, P.			
TITLE Direct Submission			
JOURNAL Submitted (13-FEB-1998) Sperling P., Physiologie, Institut fuer			
Allgemeine Botanik der Universitaet Hamburg, Ohnhorststrasse 18,			
D-22609 Hamburg, FRG			
COMMENT Related sequence: X87143			
FEATURES Location/Qualifiers			
source 1..1610			
/organism="Brassica napus"			
/cultivar="Drakkar"			
/db_xref="taxon:3708"			
/dev_stage="18 to 35 DAF"			
/tissue_type="ripening embryos"			
5'UTR 1..50			
gene 51..1400			

Db	282	A-T--CTTAAAGATTACTCTGTTTCTGAGGTTCTTAAAGATTATAGAAAGCTTGTGTTG	338
Qy	286	ACTACTCTCAAGAGCTTCGAAGTGTCCGAGATCTCCAAGGACTACCGGAGCGTTTGAACG	345
Db	339	AGTTTCTTAAATGGGTTGTATGACAAAAGGTTCATATTATCTTTTGCAACTTTGTGCT	398
Qy	346	AGATGTCCGGTCCGGGATCTTCGAGAGAAGAGGCCACCACATCATGTGACAGTTCGTGC	405
Db	399	TTATAGCAATGCTGTTTGCTATGAGTGTATTATGGGGTTTCTTTTGTGAGGGTGTTTTGG	458
Qy	406	CGGTTGGGTCATGATGGCGCAATCGTCTACGGCGTCTCGCGTCGGAGTCCGTCCGGAG	465
Db	459	TACATTGTTTCTGGGTGTTTGATGGGGTTCCTTTGGATTCAGATGGTTCAGATTGGAC	518
Qy	466	TTCACATGCTCTCGCGCGCACGTCTGGGCTTATGTGAGCCAAAGCCGCTATGTGGGCC	525
Db	519	ATGATGTGGGCATTATATGTAGTGTCTGATTCAAGGCTTAAATAGTTTATGGGTATTT	578
Qy	526	ATGACTCCGGCCATTACCAGGTGATGCCAACCCGTGGATACACAGAAATCACGCACTCA	585
Db	579	TTGCTGCAAAATGCTTTTCAGGAATAAGTATTTGGTTGGTGGAAATGGAACCAATAATGCAC	638
Qy	586	TAGCAGCGAACATCTCTAACCGGAATCAGCATCGGGTGTGGAAGTGGACCAACAACGCC	645
Db	639	ATCACATTGCCCTGTAATAGCCCTTGAATGATGCCCTGATTTACATATATATACCATTCCTTG	698
Qy	646	ACCACCTCGCCTGCAACAGCCTCGACTACGACCCCGACCTCCAGCAGATCCCCGTAATCG	705
Db	699	TTGTGTTCTCCAAGTTTTTTTGGTTCCACTCACCTCTCATTTCTATGAGAAAGGTGTGACTT	758
Qy	706	CCGTCCTCCACCGACTCTTCAACTCCATCACCTCGGTCCTCTATGGCCGAGTCTGAAAT	765
Db	759	TTGACTCTTTATCAAGATTCTTTGTAAGTATATCAACATTTGACATTTTACCCATTAATGT	818
Qy	766	TCGACGAAAGTGGCACGGTTCCTTAGTCAGCTACACGACTGGACCTACTACCCGTCATGA	825
Db	819	GTGCTGTAGGCTCAATATGTATGTACAACTCTCATAAATGTTCTTGACCAAGAAATG	878
Qy	826	TCCTTGGCCGAGTCACCTCTTCATCCAGACCTTTTATGTCTCTCACCAGGCGCGACG	885
Db	879	TGTCCTATCGAGCTCAGGAACCTCTTGGATGCCTAGTGTCTCGATTGTGTACCCGTTCG	938
Qy	886	TCCTTGACCGGCTCTAAACTTAAATGGGTATCGGGTTTTCTGGACGTTGGTTCCTGCT	945
Db	939	TTGTTCTTGTGTTGCTTAATTTGGGTGAAAGAAATATGTTTGTATTGCAAGTTTATFCAG	998
Qy	946	TCGTATCTTGTCTCCGGAACCTGGCCTGAACGGTTCGGGTTGTCCTCATCAGCTTTTCGG	1005
Db	999	TGACTGGAATGCAACAAGTTCAGTTCTCCTTGAAACCACTCTCTTCAAGTGTATTATGTTG	1058
Qy	1006	TCAGGGGATCCAGCAGCTCCAGTTTCAGCTCAACCACTTCTCCGGCGCACATACGTGG	1065
Db	1059	GAAAGCCTAAAGGGAATAATTGGTTTGAGAACAAACGGATGGGACACTTTGACATTTCTTT	1118
Qy	1066	GCCTCCCAAGGCGCAACTGGTTCGAGAAGCAGACGAAAGGACGATCGATATACGT	1125
Db	1119	GTCTCTTGTGGATGGATTGGTTTCATGTTGATGGATTCGAATTCCAATTTGAGGATCATTTGT	1178
Qy	1126	GCCACCGTGGATGGACTGGTTCTTTTGTGGGCTGCAGTTCCTCAGTTGGAGCCACACTTGT	1185
Db	1179	TTCCCAAGATGCTAGATCAACCTTAGGAAATCTCCGCCCTAGCTGATCGAGTTATGCA	1238
Qy	1186	TCCCTAGGCTGCCGCTGGCGAGCTTAGGAAGATTCGCCCTTGGCTCGGGACTTGTGTA	1245
Db	1239	AGAAACA 1245	
Qy	1246	AGAACA 1252	

RESULT	6			
LOCUS	138430	1685 bp	DNA	PAT
DEFINITION	Sequence 4 from patent US 5614393.			
ACCESSION	138430			

[illegible]

LOCUS AR024229 965 bp DNA PAT 04-DEC-1998
DEFINITION Sequence 22 from patent US 5795961.
ACCESSION AR024229
NID g3977523
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 965)
AUTHORS Wallace,T.Paul, Harris,W.J., Carr,F.J., Old,L.J., Welt,S. and Kitamura,K.
TITLE Recombinant human anti-Lewis b antibodies
JOURNAL Patent: US 5795961-A 22 18-AUG-1998;
FEATURES Location/Qualifiers
source 1..965
BASE COUNT 192 a 170 c 226 g 205 t 172 others
ORIGIN

Query Match 2.7%; Score 46; DB 22; Length 965;
Best Local Similarity 17.4%; Pred. No. 7.60e-12;
Matches 26; Conservative 76; Mismatches 44; Indels 3; Gaps 3;

Db 786 VHYSGGVSTCTASDYYTSYCWVGRGWGDYGGYTNVNGRGRVTMADTSSNRSS 845
QY 104 GTCGGCGATCTCTGGATCTCCATCCAGGCAAGGTCTACGACTCTCTCGTGGCGGC 163
Db 846 VTAADTAVYCVGRSVDSDGWDYGGTTVTVSSHTVKDMTSSSSASVGDRTVTCRSST 905
QY 164 GGA-GCACCCCGGGGAGGTCCCGTCTCCTCAGTCTGGC-CGGC-CAGGACGTACCGCA 220
Db 906 HGNNTYVYKAKYRVSNRSGVSRGS 934
QY 221 CGCTTCATTGCTACCAACCGGCGCACGG 249

RESULT 10
LOCUS RYENRA 281 bp DNA PLN 05-APR-1996
DEFINITION Secale cereale nitrate reductase gene, partial cds.
ACCESSION L40154
NID 9685205
KEYWORDS nitrate reductase.
SOURCE Secale cereale DNA.
ORGANISM Secale cereale
Eukaryotes; mitochondrial eukaryotes; Viridiplantae;
Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
Liliopsida; Poales; Poaceae; Secale.
REFERENCE 1 (bases 1 to 281)
AUTHORS Zhou,J., Kilian,A., Warner,R.L. and Kleinhofs,A.
TITLE Variation of nitrate reductase genes in selected grass species
JOURNAL Genome 38 (5), 919-927 (1995)
MEDLINE 96076701
FEATURES Location/Qualifiers
source 1..281
/organism="Secale cereale"
/db_xref="taxon:4550"
CDS <1..>281
/codon_start=1
/product="nitrate reductase"
/db_xref="PID:g685206"
/translation="DDDKQFTMSEVRKHGSKSAWIVVHGVHVDCTGFLKDHDPGGADS
ILINAGTCTEEFDATHSDKAKALLDAHRIGELITTGAGYNSDNSVHAAP"

BASE COUNT 60 a 101 c 75 g 45 t
ORIGIN

Query Match 2.5%; Score 43; DB 24; Length 281;
Best Local Similarity 63.0%; Pred. No. 6.13e-10;
Matches 104; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Db 61 TGGATCTCTCCACGGACACGTCTACGACTGACCGGCTTCTCAAGGACACCCCGGT 120
QY 117 TGGATCTCCATCCAGGGCAAGGTCTACGACTGCTCTCGTGGCGGCGGACACCCCGGC 176

LOCUS 166494 7218 bp DNA PAT 23-DEC-1997
DEFINITION Sequence 14 from patent US 5670367.
ACCESSION I66494
NID g2724471
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 7218)
AUTHORS Dörner,F., Scheiflinger,F. and Falkner,F.Gunter.
TITLE Recombinant fowlpox virus
JOURNAL Patent: US 5670367-A 14 23-SEP-1997;
FEATURES Location/Qualifiers
source 1..7218
/organism="unknown"
BASE COUNT 1944 a 1491 c 1486 g 1929 t 368 others
ORIGIN

Query Match 5.3%; Score 91; DB 22; Length 7218;
Best Local Similarity 0.3%; Pred. No. 2.20e-43;
Matches 1; Conservative 229; Mismatches 139; Indels 0; Gaps 0;

Db 1067 YY 1126
QY 634 CCACAAGCGCCACCACCTCGCTGCAACAGCTCGACTAGCAGCCCGACCTCCAGCACA 693
Db 1127 YY 1186
QY 694 TCCCGTATTGCGCTCCACCGGACTCTCAACTCCATCCTCGTCTTCTATGGCC 753
Db 1187 YY 1246
QY 754 GAGTCTGAATTCGAGAGTGGCAGGTCTTAGTACGACTACCACTGACCTACT 813
Db 1247 YY 1306
QY 814 ACCGGTCATGATCTCGGCGAGTCAACCTCTCATCAGACCTTTTATGCTCTCA 873
Db 1307 YY 1366
QY 874 CCAGCGCAGCTCCCTGACCGGCTCTAAACTAATGGGTATCGCGGTTTCTGGACGT 933
Db 1367 YY 1426
QY 934 GGTTCGGCTCTTCGATCTGTCTCCGAACTGGCTGAACGGTTCGGGTCTGCTCTCA 993
Db 1427 YYYYYYTG 1435
QY 994 TCAGCTTG 1002

RESULT 9

LOCUS AR024229 965 bp DNA PAT 04-DEC-1998
DEFINITION Sequence 22 from patent US 5795961.
ACCESSION AR024229
NID g3977523
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 965)
AUTHORS Wallace,T.Paul, Harris,W.J., Carr,F.J., Old,L.J., Welt,S. and Kitamura,K.
TITLE Recombinant human anti-Lewis b antibodies
JOURNAL Patent: US 5795961-A 22 18-AUG-1998;
FEATURES Location/Qualifiers
source 1..965
BASE COUNT 192 a 170 c 226 g 205 t 172 others
ORIGIN

Query Match 2.7%; Score 46; DB 22; Length 965;
Best Local Similarity 17.4%; Pred. No. 7.60e-12;
Matches 26; Conservative 76; Mismatches 44; Indels 3; Gaps 3;

Db 786 VHYSGGVSTCTASDYYTSYCWVGRGWGDYGGYTNVNGRGRVTMADTSSNRSS 845
QY 104 GTCGGCGATCTCTGGATCTCCATCCAGGCAAGGTCTACGACTCTCTCGTGGCGGC 163
Db 846 VTAADTAVYCVGRSVDSDGWDYGGTTVTVSSHTVKDMTSSSSASVGDRTVTCRSST 905
QY 164 GGA-GCACCCCGGGGAGGTCCCGTCTCCTCAGTCTGGC-CGGC-CAGGACGTACCGCA 220
Db 906 HGNNTYVYKAKYRVSNRSGVSRGS 934
QY 221 CGCTTCATTGCTACCAACCGGCGCACGG 249

RESULT 10
LOCUS RYENRA 281 bp DNA PLN 05-APR-1996
DEFINITION Secale cereale nitrate reductase gene, partial cds.
ACCESSION L40154
NID 9685205
KEYWORDS nitrate reductase.
SOURCE Secale cereale DNA.
ORGANISM Secale cereale
Eukaryotes; mitochondrial eukaryotes; Viridiplantae;
Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
Liliopsida; Poales; Poaceae; Secale.
REFERENCE 1 (bases 1 to 281)
AUTHORS Zhou,J., Kilian,A., Warner,R.L. and Kleinhofs,A.
TITLE Variation of nitrate reductase genes in selected grass species
JOURNAL Genome 38 (5), 919-927 (1995)
MEDLINE 96076701
FEATURES Location/Qualifiers
source 1..281
/organism="Secale cereale"
/db_xref="taxon:4550"
CDS <1..>281
/codon_start=1
/product="nitrate reductase"
/db_xref="PID:g685206"
/translation="DDDKQFTMSEVRKHGSKSAWIVVHGVHVDCTGFLKDHDPGGADS
ILINAGTCTEEFDATHSDKAKALLDAHRIGELITTGAGYNSDNSVHAAP"

BASE COUNT 60 a 101 c 75 g 45 t
ORIGIN

Query Match 2.5%; Score 43; DB 24; Length 281;
Best Local Similarity 63.0%; Pred. No. 6.13e-10;
Matches 104; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Db 61 TGGATCTCTCCACGGACACGTCTACGACTGACCGGCTTCTCAAGGACACCCCGGT 120
QY 117 TGGATCTCCATCCAGGGCAAGGTCTACGACTGCTCTCGTGGCGGCGGACACCCCGGC 176

Cp	738	AGGTGATGAGTTGAAGAGTCGGTGGACAGCGCAATACGGGGATGTCTGGAGGTCGG	679
Db	896	RVTCSRSTTHGNNGTYWYKRGAKIRVSNRSGVSRSOGSGTDYTSSDATYYC	950
Cp	678	GGT-CGTAGTCGAGCTGTTCAGGCAGAGTGCTGGCGTGTCGGTCCACTTCC	625
RESULT	13		
LOCUS	AF012089	10772 bp	DNA
DEFINITION	Drosophila melanogaster cysteine proteinase-1 (CP1) gene, complete cds, and phenylalanyl tRNA synthetase gene, partial cds.		
ACCESSION	AF012089		
NID	92305220		
SOURCE	fruit fly.		
ORGANISM	Drosophila melanogaster Eukaryota; Eukaryotic; Insecta; Pterygota; Diptera; Brachycera; Tracheata; Hexapoda; Insecta; Pterygota; Diptera; Brachycera; Muscomorpha; Epiphyrioidea; Drosophilidae; Drosophila. 1 (bases 4546 to 4553) Gray, Y.H., Tanaka, M.M. and Sved, J.A. P-element-induced recombination in Drosophila melanogaster: hybrid element insertion Genetics 144 (4), 1601-1610 (1996)		
JOURNAL			
MEDLINE	97132596		
REFERENCE	2 (bases 1 to 10772) Gray, Y.H.M., Sved, J.A., Preston, C.R. and Engels, W.R. Structure of the cysteine proteinase (CP1) gene of Drosophila melanogaster and associated mutational effects Unpublished		
JOURNAL			
REFERENCE	3 (bases 1 to 10772) Gray, Y.H.M., Sved, J.A., Preston, C.R. and Engels, W.R. Direct Submission Submitted (30-JUN-1997) School of Biological Sciences, University of Sydney, Biology A12, Sydney University, NSW 2006, Australia		
TITLE	Location/Qualifiers		
FEATURES	Source		
mRNA	1..10772 /organism="Drosophila melanogaster" /db_xref="taxon:7227" join(872)..1000,2310..2426,6476..6690,6751..7707) /gene="CP1"		
gene	/product="cysteine protease" 872..7707 /gene="CP1"		
exon	872..1000 /gene="CP1"		
intron	number=1 1001..2309 /gene="CP1"		
exon	2310..2426 /gene="CP1"		
CDS	number=2 join(2328..2426,6476..6690,6751..7462) /gene="CP1" /codon_start=1 /product="cysteine proteinase-1" /db_xref="PID:g2305221" translation="MRTAVILPLLLAVAAQVFADVMEEHHTFKLEHRKNYQDET EERELRIENKKIKAHQRRAEKVSFLKLVNKIADLHEFRLQNLNGLNYLHK QLRADEFGKGVTFISPAHVTLPRVDWRKTGAVTKDGCHGSQWASSTGALEQG HFRKSGLVLSLSEQNLDVCSITKNGNCNGMDNAPRYIKDNGIDTBKSPYPEAID DSCHFNGTAVDRGTDPDQGEKKMAEAVTVGPSVAIDASHESTQFYSEGYN EPQDAONLDHGVLVVGFDTDESEDYWLVKNSWGTTGDKGIKMLRNKENOQGIAS ASSVPLV"		
intron	2427..6475 /gene="CP1"		
misc_feature	4546..4553 /gene="CP1"		
exon	/note="insertion site of P(CaSpeR)(50C)" 6476..6690 /citation=[1] /gene="CP1"		
	number=3		

[illegible]

repeat_region	3707. .3728 /note="(A)22" /rpt_type=tandem /rpt_unit=A 4366. .4661 /rpt_family="Alu" 5327. .5602 /rpt_family="Alu" 6586. .6956 /rpt_family="L1" 6647. .6684 /note="(CA)19" /rpt_type=tandem /rpt_unit=CA 7113. .7373 /rpt_family="Alu" complement(7830. .8185) /standard_name="possible repeat" 8238. .8503 /rpt_family="Alu" 9070. .9387 /rpt_family="Alu" complement(9740. .9845) /rpt_family="MER42" complement(10440. .11015) /rpt_family="Alu" 11950. .12350 /rpt_family="Alu" 12057. .12085 /note="(A)29" /rpt_type=tandem /rpt_unit=A 12365. .12645 /rpt_family="Alu" 13727. .13750 /note="(AC)12" /rpt_type=tandem /rpt_unit=AC 13783. .14024 /rpt_family="L1" 14175. .14470 /rpt_family="Alu" complement(14906. .15259) /standard_name="possible repeat" 15300. .15613 /rpt_family="Alu" 16871. .16890 /note="(T)20" /rpt_type=tandem /rpt_unit=T complement(16675. .16977) /rpt_family="Alu" complement(16993. .17085) /rpt_family="MER42" complement(17678. .18276) /rpt_family="Alu" 19305. .19583 /rpt_family="Alu" 19914. .19945 /note="(GTTT)8" /rpt_type=tandem /rpt_unit=GTTT complement(19943. .20222) /rpt_family="Alu" 21202. .21496 /rpt_family="Alu" 21736. .22035 /rpt_family="Alu" 22017. .22038 /note="(A)22" /rpt_type=tandem /rpt_unit=A 22321. .22457 repeat_region
---------------	--

```
repeat_region /rpt_family="Alu"
repeat_region complement(22638..22981)
repeat_region /standard_name="possible repeat"
repeat_region 23022..23326
repeat_region /rpt_family="Alu"
repeat_region 23473..23761
repeat_region /rpt_family="Alu"
repeat_region 23744..23767
repeat_region /note="(A)24"
repeat_region /rpt_type=tandem
repeat_region /rpt_unit=A
repeat_region complement(24639..24694)
repeat_region /rpt_family="MER42"
repeat_region complement(25349..25713)
repeat_region /rpt_family="Alu"
repeat_region complement(25727..26471)
repeat_region /rpt_family="Alu"
repeat_region 27191..27477
repeat_region /rpt_family="Alu"
repeat_region 27774..28057
repeat_region /rpt_family="Alu"
repeat_region 28040..28066
repeat_region /note="(A)27"
repeat_region /rpt_type=tandem
repeat_region /rpt_unit=A
repeat_region 28712..28930
repeat_region /rpt_family="MER20"
repeat_region complement(28769..28838)
repeat_region /note="GRAIL 2 excellent exon, frame 0"
repeat_region 28987..29214
repeat_region /rpt_family="Alu"
repeat_region join(29485..29559,29598..29779)
repeat_region /note="95% identity dbSTS:G14522 (SHGC-11312)"
repeat_region 29495..29976
repeat_region /standard_name="AI027942"
repeat_region /note="100% identity EST ov84a10.x1"
repeat_region complement(30401..30536)
repeat_region /note="GRAIL 2 excellent exon, frame 2"
repeat_region complement(30682..30733)
repeat_region /rpt_family="MIR"
repeat_region complement(31573..31724)
repeat_region /note="GRAIL 2 excellent exon, frame 1"
repeat_region complement(32159..32232)
repeat_region /note="GRAIL 2 excellent exon, frame 2"
repeat_region 32388..32488
repeat_region /rpt_family="MLT1"
repeat_region 32617..32908
repeat_region /rpt_family="Alu"
repeat_region 32977..33088
repeat_region /rpt_family="MLT1"
repeat_region complement(33670..33785)
repeat_region /rpt_family="Alu"
repeat_region complement(34021..34144)
repeat_region /note="GRAIL 2 excellent exon, frame 2"
repeat_region complement(35238..35331)
repeat_region /note="GRAIL 2 excellent exon, frame 1"
repeat_region 36392..36663
repeat_region /rpt_family="Alu"
repeat_region 36901..37222
repeat_region /note="100% identity EST ou55c09.x1"
repeat_region /db_xref="dbEST:AI025011"
repeat_region 36901..37164
repeat_region /standard_name="STSG-9983"
repeat_region /db_xref="dbSTS:G26554"
repeat_region complement(37595..37654)
repeat_region /note="GRAIL 2 excellent exon, frame 0"
repeat_region complement(join(38063..38218,38462..38578,38741..38995,
repeat_region 39071..39205,39532..39630,39935..40048,40300..40410,
repeat_region 40503..40561,41868..41972,42103..42225,42492..42569,
repeat_region 44379..44507))
repeat_region /standard_name="histidyl-tRNA synthetase"
repeat_region /note="78%-100% protein identity GenPept:U18937"
repeat_region complement(38069..38215)
```

```
misc_feature /note="GRAIL 2 excellent exon, frame 0"
... complement(38462..38578)
Note: remainder of annotations omitted.
Query Match 2.2%; Score 38; DB 27; Length 74371;
Best Local Similarity 13.3%; Pred.No. 7.36e-07;
Matches 28; Conservative 101; Mismatches 79; Indels 2; Gaps 2;
Db 15925 SKSCMSRSKRGWYRSWKYRCAMWMTCKSSKRCWCWSYRMKCYSCSYCYS-SGK 15983
QY 774 GTGGCAGGTTCCCTAGT-CAGCTACCAGCAGCTGGACCTACTACCGGGTCATGATCTTCGG 832
Db 15984 KYWRCRSMYWTTCYYSKYYSMSYCTCTSMGWRWMMWSKGRSMWYASRSGCSCSMCM 16043
QY 833 CCGAGTCAACCTCTTCATCCAGACCTTTTATTGCTCCTCACCAGGCGGACGTCCTGA 892
Db 16044 CRCSCMSMKMWWTTTNTTKRTTWTTWKRRKAGASRGSKRCRCMSYGGKSTCKMKMTC 16103
QY 893 CCGCGCTCTAAACTTAATGGGTATCGCGGTTTCTGAGAGTGGTTCGCGCTCTTCGTATC 952
Db 16104 YYGMYSWCSYRMKYSRYCCGCYSSSYCKS 16133
QY 953 TTGCTCCCGAAGTGGCTGAACGGTTCGG 982
```

Search completed: Thu Feb 18 16:22:21 1999
Job time : 6092 secs.

W P E R L H (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Thu Feb 18 17:40:33 1999; Maspar time 244.06 Seconds
948.786 Million cell updates/sec

Local output not generated.

Title: >US-08-934-254-26
Description: (1-1702) from US08934254.seq

Perfect Score: 1702
N.A. Sequence: 1 CCCCAAAATTTTCATGTT.....TTTTGGTAAAAAAAAA 1702
Comp: GGGGTTTTTAAAGTAACAA.....AAAAACCATTTTTTTTTT

Scoring table: TABLE default

Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 188442 seqs, 68026449 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: n-geneseq32

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39 40:part40

Statistics: Mean 9.489; Variance 5.740; scale 1.653

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description	Pred. No.
1	172	10.1	T30395	Borage delta-6-desatu	2.48e-94
2	47	2.8	204	Base substituted E.co	8.70e-13
3	43	2.5	91	Oligonucleotide probe	1.79e-10
4	43	2.5	204	Base substituted E.co	1.79e-10
5	40	2.4	91	Oligonucleotide probe	9.00e-09
6	36	2.1	114	Generic DNA sequence	1.47e-06
7	36	2.1	114	Generic DNA sequence	1.47e-06
8	36	2.1	114	Generic DNA sequence	1.47e-06
9	36	2.1	114	Generic DNA sequence	1.47e-06
10	36	2.1	114	Generic DNA sequence	1.47e-06
11	35	2.1	114	Generic DNA sequence	5.12e-06
12	36	2.1	114	Generic DNA sequence	1.47e-06
13	36	2.1	114	Generic DNA sequence	1.47e-06

C 14	36	2.1	114	12	070467	Generic DNA sequence	1.47e-06
C 15	35	2.1	114	12	070469	Generic DNA sequence	5.12e-06
C 16	34	2.0	114	12	070466	Generic DNA sequence	1.77e-05
C 17	34	2.0	114	12	070471	Generic DNA sequence	1.77e-05
C 18	32	1.9	114	12	070473	Generic DNA sequence	2.02e-04
C 19	32	1.9	114	12	070472	Generic DNA sequence	2.02e-04
C 20	32	1.9	114	12	070466	Generic DNA sequence	2.02e-04
C 21	32	1.9	114	12	070470	Generic DNA sequence	2.02e-04
C 22	32	1.9	178	32	T76405	Human endothelin-1 an	2.02e-04
C 23	30	1.8	172	32	T76363	Human interleukin 8 a	2.19e-03
C 24	30	1.8	178	32	T76405	Human endothelin-1 an	2.19e-03
C 25	30	1.8	250	32	T76438	Substance P antisense	2.19e-03
C 26	31	1.8	540	32	T75438	Mortierella alpina cy	6.70e-04
C 27	30	1.8	565	6	Q35072	HCV envelope region n	2.19e-03
C 28	29	1.7	172	32	T76363	Human interleukin 8 a	7.04e-03
C 29	29	1.7	190	32	T76452	Chymase antisense oli	7.04e-03
C 30	29	1.7	190	32	T76452	Chymase antisense oli	7.04e-03
C 31	27	1.6	39	7	Q51787	Mixed oligonucleotide	6.94e-02
C 32	27	1.6	67	24	T14322	Primer used in the la	6.94e-02
C 33	28	1.6	114	12	070471	Generic DNA sequence	2.23e-02
C 34	28	1.6	114	12	070473	Generic DNA sequence	2.23e-02
C 35	28	1.6	130	32	T76152	Human vascular cell a	2.23e-02
C 36	27	1.6	168	32	T76270	Human MDNCF antisense	6.94e-02
C 37	26	1.5	90	21	T13616	ME#1 random peptide 1	2.12e-01
C 38	26	1.5	201	2	N70195	Streptomyces protease	2.12e-01
C 39	26	1.5	201	2	N70194	Signal portion of gen	2.12e-01
C 40	26	1.5	1558	22	T28255	Survival motor neuron	2.12e-01
C 41	26	1.5	1560	22	T18828	Human survival motor	2.12e-01
C 42	26	1.5	1582	22	T18831	Human survival motor	2.12e-01
C 43	26	1.5	1582	22	T28259	Survival motor neuron	2.12e-01
C 44	26	1.5	3871	2	N71302	HSV-1 gB and surround	2.12e-01
C 45	26	1.5	5852	2	Q11710	Dictyostellium plasmid	2.12e-01

ALIGNMENTS

RESULT 1

ID T30395 standard; DNA; 1685 BP.
AC T30395;
DE 15-SEP-1996 (first entry)
DT Borage delta-6-desaturase gene.
KW Delta-6-desaturase; gamma-linolenic acid; transgenic plant; borage;
KW polyunsaturated fatty acid; octadecatetraenoic acid;
KW chilling resistance; oilseed; ss; ds.
OS Borage officinalis.
PH Key Location/Qualifiers
FT cds 44..1390 /*tag= a
FT W09621022-A2.
PN 11-JUL-1996.
PD 28-DEC-1995; IB1167.
PF 30-DEC-1994; US-366779.
PR (RHON) RHONE POULENC AGROCHIMIE.
PA Freyssonnet GL, Nuccio M, Nunberg AN, Reddy AS, Thomas TL;
PI WPI: 96-333997/33.
DR P-PSDB: R98455.
DR Transgenic plants comprising the borage delta-6-desaturase gene
PT show increased production of gamma linolenic acid and having
PT increased resistance to chilling
PS Claim 2; Page 51-52; 75pp; English.
CC A DNA clone (T30395) codes for borage delta-6-desaturase (R98455),
CC which catalyses the conversion of linoleic acid to gamma-linolenic
CC acid (GLA). It was isolated from a borage membrane-bound polysomal
CC library using probes based on abundantly expressed seed storage
CC protein cDNAs and with an isolated partial cDNA clone. The gene
CC can be incorporated into a vector, pref. incorporating a
CC tissue-specific promoter, for the expression of delta-6-desaturase
CC in transgenic plants, esp. sunflower, soybean, maize, tobacco,
CC peanut, carrot or oilseed rape, resulting in increased GLA prodn.
CC Alteration of plant lipids may also lead to improved chilling
CC resistance.
SQ Sequence 1685 BP; 431 A; 277 C; 357 G; 620 T;

```
Query Match      10.1%; Score 172; DB 21; Length 1685;
Best Local Similarity 57.7%; Pred. No. 2,48e-94;
Matches 697; Conservative 0; Mismatches 507; Indels 3; Gaps 2;

Db 42 caatgctgctcaaatcaagaatacattaccctcagatgaactcaagaaccacgataaac 101
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 46 CAATGAGGGCGAAGCTAAGAAGTATATACCGCGGAGGACCTCCGCGCCCAACAACAGT 105
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 102 ccggagatctatgattctcagatcaaggaagagcctatgatcttccgattgggtgaaag 161
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 106 CCGGGGATCTCTGGATCTCCATCCAGGGCAAGGTCTACGACTGCTCTCGGTGGCGCGCG 165
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 162 accatccaggtgagcgtcttccttgagagctctgctcagaggtcaagaggttaactgatgat 221
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 166 AGCACCCCGCGCGGAGGTCCTCCGCTCTCAGTCTGGCCGCGCAGGACGTCACCGAGCCT 225
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 222 ttgttgatccatcctgctctacatgaagaatacttgataagtttttcaactgggtatt 281
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 226 TCATTGCGTACCAACCCGCGCACGGGTGCGGCATCTGGATCGGCTCTTCACCGGCTACT 285
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 282 a-t--ctaaagattactctgtttctgaggtttctaaagattatggaagcttgggttg 338
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 286 ACTACTCAAGGACTTCGAAGTGTGGAGATCTCCAAGGACTACCGGAGGCTTTTGAAGC 345
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 339 agtttttcaaatgggtttgtatgacaaaagaggtcatattatgtttgcaactttgtgct 398
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 346 AGATGTCGGGTCCGGGATCTTCGAGAAAGAGGGCCACCACATCATGTGGAGCTTCGTCG 405
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 399 ttatgcaatgcgtttgtgatgaggtttgtgaggtttgtttgtgaggggtgttttg 458
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 406 GGGTTGCGGTATGATGGGGCAATCGTCTACGGGGTGTGCGGTGCGAGTCCGCGAG 465
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 459 tacatttttctctgggtttgtatggggtttcttggattcagagtggttggattggac 518
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 466 TTCACATGCTCTGCGCGCACCTGCTGGGCTTCTGTGGATCTCAAGCGCGTATGTGGGCC 525
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 519 atgatctgggcatatgatgagtgctgtcattcaagccttaaatgattatgggtattt 578
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 526 ATGATCTCGGGCATTTACAGGTGATGCCAACCCGTGGATACACAGAAATCACGCCAACATCA 585
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 579 ttgctgcaattgttttcaggaataagattgtgttggaagaaatggaacataatgac 638
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 586 TAGCAGGCAACATCTACCGGAATCAGATCGCTGTTGGAAGTGGACCCACACAGGCC 645
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 639 atcacattgcttgaatgacctgattgacctgatttacaatatataacattccttg 698
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 646 ACCACTCGCTGCACAGCCTCGACTACGACCCCGACCTCCAGCACATCCCGCTATTCG 705
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 699 ttgtgttccaagtgttttgggtcactcactcctcattcttatgagaaaaggttgactt 758
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 706 CGGTCTCCACCGGACTTCAACTCCATCACCTCGGTCTCTATGCGCGAGTCCTGAAT 765
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 759 ttgactctttatcaagattcttgaattatcaacattggacattttacccattattgt 818
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 766 TCGACGAAGTGCACGGTCTCTATCCAGACCTTTTATTTGCTCTCACCGCGCGAGC 825
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 819 gtgctgtaggtcctaatgtatgtacatctctcattatgttggtagcagaagaaatg 878
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 826 TCTTCGGCGGAGTCAACCTTCTATCCAGACCTTTTATTTGCTCTCACCGCGCGAGC 885
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 879 tgtcctatcgagctcaggaactcttggatgctcagtgcttctcgatttggtagccgttgc 938
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 886 TCCCTGACCGGCTCTAAACTTAATGGGTATCGCGGTTTCTGGACGTGTTTCCCGCTCT 945
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 939 ttgttttcttggctaatgtgggtgaaagaaatattgtttgtttatttcaagtttattcag 998
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 946 TCGTATCTGTCTCCGAACCTGGCTGAACCGGTTCGGGTTCTGCTCTCATCAGCTTTCGG 1005
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 999 tgaactggaatcaacaagtctcagttctccttggaaccattcttctcaagtgtttatgtg 1058
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1006 TCAGCGGATCCAGACGCTCCAGTTCAGCTCAACCACTTCTCCGGCGACACATACGTTGG 1065
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1059. gaaagcctaaaggggaaataatgtgttgagaaacaaagcgtgggacatttgacatttctt 1118
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```


QY 624 TGGAAGTGGACCCACACACGCCCCACCCTCGCCTGCACAGCCTCGACTACGACCCCGCAC 6B

Db 94 yzrmrbnvdydnrsdaaaavcyvrrsvkvdccynachddhhvybbvbnvnvhnncnc 15

QY 684 CTCGAGCAATCCCGGTATTTCGGCGTCTCCACCGGACTCTTCAACTCATCAGCTCGTC 74

Db 154 ccbnnhvchnvbnnhrnwayvrhdarrdvh 185

QY 744 TTCTATGCCGAGTGCTGAAATTCGACGGAAGT 775

RESULT 5

ID Q51746 standard; cDNA; 91 BP.

AC Q51746;

DT 31-MAY-1994 (first entry)

DE Oligonucleotide probe MK14-A

KW Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;

OS Synthetic.

PN EP-571911-A.

PD 01-DEC-1993.

PF 24-MAY-1993; 108325.

PR 28-MAY-1992; US-889651.

PI (BECK) BECTON DICKINSON CO.

PI Shank DD, Spears PA;

PI WPI; 93-378844/48.

PT New oligo-nucleotide probes specific for Mycobacteria - used for

PT detection and amplification of Mycobacteria nucleic acid in

PS samples

PS Claim 3; Page 14; 23pp; English.

CC Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14

CC (O51735). It hybridized to all spp. of mycobacteria tested, but

CC cross reacted to a few non-mycobacterial spp. The probe may

CC be useful as an initial screen for mycobacterial infection.

CC See also Q51735-45 and Q51747-59.

SQ Sequence 91 BP; 5 A; 17 C; 15 G; 4 T;

Query Match 2.4%; Score 40; DB 9; Length 91;

Best Local Similarity 5.5%; Pred. No. 9.00e-09;

Matches 3; Conservative 47; Mismatches 4; Indels 1; Gaps

Db 10 gssvhsyvvvvhshbvhhvhhvhsvvvvhvvhv-vhhvhyhvyvsctc 63
|::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ||
Cp 1537 GGGAACGCAACCACTAGCAAAGTTAAGCAAGTAAGCTGAAGCTAGTGCTTC 1483

RESULT 6

ID Q70469 standard; DNA; 114 BP.

AC Q70469;

DT 07-APR-1995 (first entry)

DE Generic DNA sequence to generate a random TSAR peptide library.

KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;

KW effector domain; concatenated heterofunctional protein; linker;

KW direct; rapid; detection; screening; treatment; generic; ss..

OS Synthetic.

FH Key Location/Qualifiers

FT misc_feature 55..60

FT /*tag= a

FT /note= "this sequence represents 'Z'; Z can be a

FT sequence of 6,9 or 12 nucleotides (see

FT comments")

PN WO9418318-A.

PD 18-AUG-1994.

PF 01-FEB-1994; U00977.

PR 01-FEB-1993; US-013416.

PR 30-DEC-1993; US-176500.

PR 31-JAN-1994; US-189331.

PA (UNIC-) UNIV NORTH CAROLINA.

PI Fowlkes DM, Kay BK;

PI WPI; 94-279739/34.

PT Identifying proteins or peptide(s) which bind a ligand - by

PT screening a recombinant vector library expressing fusion proteins

PT comprising a binding domain and an effector domain

[illegible]

RESULT 10
ID Q70467 standard: DNA: 114 BP.

AC Q/0467;
DT 05-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.

DE generic DNA sequence to generate a random ISAK peptide library.
 KW TEAR; totally synthetic affinity reagent; synthetic; binding domain;
 KW effector domain; concatenated heterofunctional protein; linker;
 KW direct; rapid; detection; screening; treatment; generic; ss.
 OS Synthetic.

OS Synthetic.
 KW direct; rapid; detection; screening; treatment; generic; ss.
 KW effector domain; concatenated heterofunctional protein; linear
 KW effector domain; concatenated heterofunctional protein; linear

OS	Synthetic.	
FH	Key	Location/Qualifiers
FT	misc feature	55..60

```

FT      /*tag= a
FT      /note= "this sequence represents 'Z'; Z can be a
FT      sequence of 6, 9 or 12 nucleotides (see
FT      comments)"

```

FT	comments)
PN	WO9418318-A.
PD	18-AUG-1994.

PD 18-AUG-1994. U00977.
 PF 01-FEB-1994; U00977.
 PP 01-FEB-1993. US-013416

PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 21-TAM-1994; USC-180321.
PR 21-TAM-1994; USC-180321.

PR 31-JAN-1994; US-189331.
PA (UYNCC-) UNIV NORTH CAROLINA.
BT

PI Fowlkes DM, Kay BK;
DR WPI; 94-279739/34.

PT Identifying proteins or peptide(s) which bind a ligand - by

PT screening a recombinant vector library expressing fusion protein comprising a binding domain and an effector domain

CC Q70467 is a generic DNA sequence used to generate random TSAR (T

CC Synthetic Affinity Reagents) peptides. This generic formula can be
CC represented as follows: X(NNB)₁(TGC)₁(NNB)₁(TGC)₁(NNB)₁LY
CC Q/U46/ is a generic DNA sequence used to generate random TSAE

CC represented as follows: X(NNB)16(TGC)(NNB)1Z(NNB)16(TGC)(NNB)1LY
CC and Y are flanking restriction sites (X is not the same as Y) t

CC not specified further. Other generic sequences are shown in Q70404 and a defining accession (Q70404) is not the same as Q70404. Other generic sequences are shown in Q70404. Other specific peptides generated by these generic sequences are shown in Q70404.

CC comprising at least two functional regions - a binding domain wi
CC R65151-54. TSARs are concatenated heterofunctional proteins or
CC other specific peptides generated by these generic sequences an

CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that
CC chemically or biologically active. They may further comprise a li

CC peptide between the 2 domains. The oligonucleotides are also des-
CC that the expressed peptide contains 2 or 4 cysteine residues posi-
CC tively. The oligonucleotides are also des-
CC chemically or biologically active. They may further comprise a
CC peptide between the 2 domains. The oligonucleotides are also des-
CC that the expressed peptide contains 2 or 4 cysteine residues posi-

CC in, or flanking, the unpredicted or variant residues. These resi-

CC or confers a TSAR binding domain can be used in vivo
CC confer some degree of conformational rigidity to the peptides.
CC deliver a chemically or biologically active moiety or metal ion

CC deliver a chemically or biologically active moiety, eg. metal
CC radioisotope, peptide, toxin or enzyme, to the specific target or
CC cell where it can also replace the function of macromolecules or

cell. They can also replace the function of macromolecules, eg. monoclonal or polyclonal antibodies and therefore circumvent the molecular method of hybridoma formation. One serious drawback of the molecular method of hybridoma formation is the time involved in the

CC complex methods of hybridoma formation or in vivo antibody produ
CC The TSARs are easily characterised and have designed activity al

CC direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 2.1%; Score 36; DB 12; Length 114;

Query Match 2.1%;
 Best Local Similarity 3.6%;
 Pred. No. 1.47e-06;
 Score 30; DB 12; Length 114;
 Matches 4; Conservative 34; Mismatches 74; Indels 0; Gaps 0

[illegible]

```

DB      3   bbbbbbcccccccccccccccccccccccccccccccccccccccccccccccccccccc
QY     71   TATCAGCGGCGGAGGACCTCCGCCGCACAACAAGTCCGGCGACTCTGTGGATCCCATCCC

```

QY 71 TATCACGGCGGAGGACCTCCGGCCGCCACAACAAAGTCCGGCGATCTCTGGATCTCCATCC

KW	effector domain; concatenated heterofunctional protein; linker;
KW	Synthetic.
FH	Key
FT	misc_feature
FT	Location/Qualifiers
FT	/tag= a
FT	/note= "this sequence represents 'Z'; Z can be a
FT	sequence of 6, 9 or 12 nucleotides (see
FT	comments)"
WT	WO9418318-A.
PN	18-AUG-1994.
PD	01-FEB-1994; UO0977.
OS	01-FEB-1994; US-013416.
PR	01-FEB-1993; US-013416.
PR	30-DEC-1993; US-176500.
PR	31-JAN-1994; US-189331.
PA	(UNCL-) UNIV NORTH CAROLINA.
PI	Fowlkes DM, Kay BK;
PB	WPI: 94-279739/34.
DR	P-PDSB; K65150 and R65151.
PT	identifying proteins or peptide(s) which bind a ligand - by
PT	screening a recombinant vector library expressing fusion proteins
PT	comprising a binding domain and an effector domain
PS	Disclosure; Page 35: 255pp; English.
CC	Q70455 is a generic DNA sequence used to generate random TSAR (Totally
CC	Synthetic Affinity Reagents) peptides.This generic formula can also be
CC	represented as follows: X(NNB)Y6(TGC)(NNB)IIZ(NNB)I4(TGC)(NNB)3Y. X
CC	and Y are flanking restriction sites (X is not the same as Y) that are
CC	not specified further. Other generic sequences are shown in Q70466-68.
CC	Other specific peptides generated by these generic sequences are shown in
CC	R65151-54. TSARS are concatenated heterofunctional proteins or peptides,
CC	comprising at least two functional regions - a binding domain with
CC	affinity for a ligand and a second effector peptide portion that is
CC	chemically or biologically active.They may further comprise a linker
CC	peptide between the 2 domains.The oligonucleotides are also designed so
CC	that the expressed peptide contains 2 or 4 cysteine residues positioned
CC	in, or flanking, the unpredicted or variant residues. These residues
CC	confer some degree of conformational rigidity to the peptides. The TSARS
CC	or compans. comprising a TSAR binding domain can be used in vivo to
CC	deliver a chemically or biologically active moiety, eg. metal ion,
CC	radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC	cell. They can also replace the function of macromolecules, eg.
CC	monoclonal or polyclonal antibodies and therefore circumvent the need
CC	for complex methods of hybridoma formation or in vivo antibody
CC	production. The TSARS are easily characterised and have designed
CC	activity allowing direct and rapid detection in a screening process.
CC	Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;
SQ	
Query Match 2.1%; Score 36; DB 12; Length 114;	
Best Local Similarity 4.7%; Pred.No.1.47e-06;	
Matches 5; Conservative 32; Mismatches 69; Indels 0; Gaps 0;	
Dd	9 bnnbnbnbtgcnnbn 68
Cp	1115 GATCGTCCCTTCGTGCTTCGGAACAGTTGCGCCCTTGGGGGGCCACGTATGT 1056
Dd	69 bnnbn 114
Cp	1055 GTCCGCGGAGAAGTGGTTGAGCGGTGAACGTGAGCGTCTCGATCGCC 1010
RESULT 13	
ID	Q70468 standard; DNA; 114 BP.
DE	AC Q70468;
DT	05-APR-1995 (first entry)
DD	Generic DNA sequence to generate a random TSAR petide library.
KW	TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW	effector domain; concatenated heterofunctional protein; linker;
KW	direct; rapid; detection; screening; treatment; generic; ss.
OS	Synthetic.
FH	Key
FT	Location/Qualifiers
FT	misc_feature
FT	55..60
FT	/tag= a
FT	/note= "this sequence represents 'Z'; Z can be a

```

sequence of 6, 9 or 12 nucleotides (see
comments)"
PT WO9418318-A.
PD 18-AUG-1994.
PN 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UINC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Key BK;
DR WPI; 94-279739/34.
DR P-PSDB; R65154.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC Q70467 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)16(TGC)(NNB)12(NNB)16(TGC)(NNB)14. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active. They may further comprise a linker
CC peptide between the 2 domains. The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or compsns. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed activity
CC allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 2.1%; Score 36; DB 12; Length 114;
Best Local Similarity 4.5%; Pred. No. 1.47e-06;
Matches 5; Conservative 33; Mismatches 74; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 62
Cp 659 GCAGGCGAGGTGGCGTGTGGTCCACTCCACCGCGATGTCGCGGTAG 600
Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnn 114
599 GATGTTGCTGCTATGATGTTGCGTGATCTGTTGATCCACGGGTTGGCATC 548

RESULT 14
ID Q70467 standard; DNA; 114 BP.
AC Q70467;
DE 05-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "this sequence represents 'Z'; Z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UINC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Key BK;
DR WPI; 94-279739/34.
DR P-PSDB; R65154.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC Q70467 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)11(TGC)(NNB)62(NNB)7(TGC)(NNB)10Y. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active. They may further comprise a linker
CC peptide between the 2 domains. The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or compsns. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed activity
CC allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 2.1%; Score 36; DB 12; Length 114;
Best Local Similarity 4.5%; Pred. No. 1.47e-06;
Matches 5; Conservative 33; Mismatches 74; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 62
Cp 659 GCAGGCGAGGTGGCGTGTGGTCCACTCCACCGCGATGTCGCGGTAG 600
Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnn 114
599 GATGTTGCTGCTATGATGTTGCGTGATCTGTTGATCCACGGGTTGGCATC 548

RESULT 14
ID Q70467 standard; DNA; 114 BP.
AC Q70467;
DE 05-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "this sequence represents 'Z'; Z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UINC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Key BK;
DR WPI; 94-279739/34.
DR P-PSDB; R65154.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC Q70467 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)16(TGC)(NNB)12(NNB)16(TGC)(NNB)14. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active. They may further comprise a linker
CC peptide between the 2 domains. The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or compsns. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed activity
CC allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 2.1%; Score 36; DB 12; Length 114;
Best Local Similarity 2.8%; Pred. No. 1.47e-06;
Matches 3; Conservative 33; Mismatches 70; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 62
Cp 737 GGTGATGAGTTGAAGAGTCGGGTGGAGACGGCGAATACGGGATGCTGGAGTCGGG 678
Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnn 108
Cp 677 GTCGTAGTCGAGGTGTTGTCAGGCGAGGTGTCGGCGTGTGGGTC 632

RESULT 15
ID Q70469 standard; DNA; 114 BP.
AC Q70469;
DE 07-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "this sequence represents 'Z'; Z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UINC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Key BK;
DR WPI; 94-279739/34.
DR P-PSDB; R65154.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC Q70467 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)16(TGC)(NNB)12(NNB)16(TGC)(NNB)14. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active. They may further comprise a linker
CC peptide between the 2 domains. The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or compsns. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed activity
CC allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

```

Search completed: Thu Feb 18 17:49:40 1999
Job time : 547 secs.

Result	Seq. No.	Query			ID	Description	Pred. No.
		Score	Match	Length			
1	220	6.3	890	1	NIA2_SOVEN	INDUCIBLE NITRATE REDU	3.17e-20
2	219	6.2	881	1	NIA1_PHAVU	NITRATE REDUCTASE 1 (E	4.87e-20
3	217	6.2	917	1	NIA1_ARATH	NITRATE REDUCTASE 1 (E	1.15e-19
4	215	6.1	911	1	NIA1_BRANA	NITRATE REDUCTASE, CLO	2.72e-19
5	212	6.0	911	1	NIA2_BRANA	NITRATE REDUCTASE, CLO	9.79e-19
6	209	5.9	136	1	CYB5_TOBAC	CYTOCHROME B5.	3.51e-18
7	203	5.8	120	1	CYB5_YEAST	CYTOCHROME B5.	4.45e-17
8	203	5.8	137	1	CYB5_ORYSA	CYTOCHROME B5.	4.45e-17
9	203	5.8	900	1	NIA1_LOTJA	NITRATE REDUCTASE (EC	4.45e-17
10	205	5.8	917	1	NIA2_ARATH	NITRATE REDUCTASE 2 (E	1.91e-17
11	202	5.7	890	1	NIA2_PHAVU	NITRATE REDUCTASE 2 (E	6.78e-17
12	200	5.7	911	1	NIA1_DUCMA	NITRATE REDUCTASE (EC	1.57e-16
13	202	5.7	918	1	NIA2_CUCMA	NITRATE REDUCTASE (EC	6.78e-16
14	199	5.7	926	1	NIA_SPIOL	NITRATE REDUCTASE (EC	2.39e-16
15	196	5.6	898	1	NIA2_BETVE	NITRATE REDUCTASE (NAD	8.35e-16
16	193	5.5	134	1	CYB5_BRAOL	CYTOCHROME B5.	2.90e-15
17	195	5.5	135	1	CYB5_TOBAC	CYTOCHROME B5, SEED IS	1.27e-15
18	195	5.5	886	1	NIA1_SOVEN	INDUCIBLE NITRATE REDU	1.27e-15
19	194	5.5	904	1	NIA1_TOBAC	NITRATE REDUCTASE 1 (E	1.92e-15
20	193	5.5	904	1	NIA2_TOBAC	NITRATE REDUCTASE 2 (E	2.90e-15
21	193	5.5	909	1	NIA2_PETHY	NITRATE REDUCTASE (EC	2.90e-15
22	192	5.5	982	1	NIA2_NEUCR	NITRATE REDUCTASE (NAD	4.39e-15
23	191	5.4	92	1	CYM5_RAT	CYTOCHROME B5, OUTER M	6.64e-15

[illegible]

CC -1- CATALYTIC ACTIVITY: NADH + NITRATE = NAD(+) + NITRITE + H(2)O.
 CC -1- COFACTOR: REQUIRES FAD, A HEME GROUP (CALLED CYTOCHROME B-557)
 CC AND ONE MOLYBDENUM ATOM.
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SIMILARITY: TO EUKARYOTIC MOLYBDOPTERIN OXIDOREDUCTASES IN THE
 CC N-TERMINAL DOMAIN.
 CC -1- SIMILARITY: CONTAINS A CYTOCHROME B5 FAMILY, HEME-BINDING DOMAIN.
 CC -1- SIMILARITY: TO FAD/NAD-BINDING CYTOCHROME REDUCTASES IN THE
 CC C-TERMINAL DOMAIN.
 DR EMBL: D38219; G540485; -
 DR PROSITE; PS00191; CYTOCHROME_B5; 1.
 DR PROSITE; PS00559; MOLYBDOPTERIN_EUK; 1.
 DR OXIDOREDUCTASE; FLAVOPROTEIN; FAD; NAD; HEME; MOLYBDENUM;
 KW NITRATE ASSIMILATION; MULTIGENE FAMILY.
 FT METAL 191 191 MOLYBDENUM-PTERIN (POTENTIAL).
 FT METAL 245 245 MOLYBDENUM-PTERIN (POTENTIAL).
 FT DISULFID 430 430 INTERCHAIN (POTENTIAL).
 FT BINDING 574 574 HEME LIGAND (BY SIMILARITY).
 FT BINDING 597 597 HEME LIGAND (BY SIMILARITY).
 SO SEQUENCE 911 AA; 102252 MW; 57A3E33B CRC32;
 Query Match 6.1%; Score 215; DB 1; Length 911;
 Best Local Similarity 38.4%; Pred. No. 2.72e-19;
 Matches 28; Conservative 18; Mismatches 26; Indels 1; Gaps 1;
 Db 535 MNTAKMYSME-VRKNSVESAWIIVHGYDCTRLKDHGGSDSLINAGTDCTEEF 593
 QY 1 MEGEAKYITAEEDLRNRKSGDLWISIQGVKDCSRWAAEHPPGVPLLSLAGQDVDAF 60
 Db 594 EAIHSDRAKALLE 606
 QY 61 IAYHPGTAWRHLD 73
 RESULT 5
 ID NTA2_BRANA STANDARD; PRT; 911 AA.
 AC P39868;
 DT 01-FEB-1995 (REL. 31, CREATED)
 DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE NITRATE REDUCTASE, CLONE PBNER1412 (EC 1.6.6.1) (NR).
 GN NTA2.
 OS BRASSICA NAPUS (RAPE).
 OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
 OC CAPPARALES; CRUCIFERAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. LISANDRA;
 RX MEDLINE; 96222419.
 RA YUKUOKA H., OGAWA T., MINAMI H., YANO H., OHKAWA Y.;
 CC -1- FUNCTION: NITRATE REDUCTASE IS A KEY ENZYME INVOLVED IN THE FIRST
 CC STEP OF NITRATE ASSIMILATION IN PLANTS, FUNGI AND BACTERIA.
 CC -1- CATALYTIC ACTIVITY: NADH + NITRATE = NAD(+) + NITRITE + H(2)O.
 CC -1- COFACTOR: REQUIRES FAD, A HEME GROUP (CALLED CYTOCHROME B-557)
 CC AND ONE MOLYBDENUM ATOM.
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SIMILARITY: TO EUKARYOTIC MOLYBDOPTERIN OXIDOREDUCTASES IN THE
 CC N-TERMINAL DOMAIN.
 CC -1- SIMILARITY: CONTAINS A CYTOCHROME B5 FAMILY, HEME-BINDING DOMAIN.
 CC -1- SIMILARITY: TO FAD/NAD-BINDING CYTOCHROME REDUCTASES IN THE
 CC C-TERMINAL DOMAIN.
 DR EMBL: D38220; G540487; -
 DR PROSITE; PS00191; CYTOCHROME_B5; 1.
 DR PROSITE; PS00559; MOLYBDOPTERIN_EUK; 1.
 KW OXIDOREDUCTASE; FLAVOPROTEIN; FAD; NAD; HEME; MOLYBDENUM;
 KW NITRATE ASSIMILATION; MULTIGENE FAMILY.
 FT METAL 191 191 MOLYBDENUM-PTERIN (POTENTIAL).
 FT METAL 245 245 MOLYBDENUM-PTERIN (POTENTIAL).
 FT DISULFID 430 430 INTERCHAIN (POTENTIAL).
 FT BINDING 574 574 HEME LIGAND (BY SIMILARITY).
 FT BINDING 597 597 HEME LIGAND (BY SIMILARITY).
 SO SEQUENCE 911 AA; 102200 MW; F0CFEF43 CRC32;

Query Match 6.0%; Score 212; DB 1; Length 911;
 Best Local Similarity 39.7%; Pred. No. 9.79e-19;
 Matches 29; Conservative 16; Mismatches 27; Indels 1; Gaps 1;
 Db 535 MNTAKMYSME-VRKNSVESAWIIVHGYDCTRLKDHGGSDSLINAGTDCTEEF 593
 QY 1 MEGEAKYITAEEDLRNRKSGDLWISIQGVKDCSRWAAEHPPGVPLLSLAGQDVDAF 60
 Db 594 EAIHSDRAKALLE 606
 QY 61 IAYHPGTAWRHLD 73
 RESULT 6
 ID CYB5_TOBAC STANDARD; PRT; 136 AA.
 AC P49098;
 DT 01-FEB-1996 (REL. 33, CREATED)
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
 DE CYTOCHROME B5.
 OS NICOTIANA TABACUM (COMMON TOBACCO).
 OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
 OC SOLANALES; SOLANACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-LEAF;
 RX MEDLINE; 94325476.
 RA SMITH M.A., STOBART A.K., SHEWRY P.R., NAPIER J.A.;
 CC -1- FUNCTION: CYTOCHROME B5 IS A MEMBRANE BOUND HEMOPROTEIN WHICH
 CC FUNCTION AS AN ELECTRON CARRIER FOR SEVERAL MEMBRANE BOUND
 CC OXYGENASES. MAY PLAY A KEY ROLE IN THE MODIFICATION BY
 CC DESATURATION OF FATTY ACIDS IN THE ENDOPLASMIC RETICULUM, WHICH IN
 CC THE DEVELOPING SEED IS UTILIZED FOR MEMBRANE SYNTHESIS AND IN THE
 CC DEVELOPMENTALLY REGULATED PRODUCTION OF LARGE AMOUNTS OF STORAGE
 CC LIPIDS. IS INVOLVED IN THE REDUCTION OF CYTOCHROME P-450 AND
 CC MAY THEREFORE BE INVOLVED IN FLAVONOID BIOSYNTHESIS IN THE PETALS.
 CC -1- SUBCELLULAR LOCATION: MICROSOMAL MEMBRANE. BOUND TO THE
 CC CYTOPLASMIC SIDE OF THE ENDOPLASMIC RETICULUM (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: IS HIGHLY EXPRESSED IN DEVELOPING SEEDS,
 CC MODERATELY EXPRESSED IN FLOWERS, AND IS EXPRESSED AT LOW
 CC LEVELS IN THE LEAF.
 DR EMBL: X71441; G296386; ALT_INIT.
 DR EMBL: X68140; G19853; -
 DR PROSITE; PS00191; CYTOCHROME_B5; 1.
 KW ELECTRON TRANSPORT; TRANSMEMBRANE; HEME; MICROSOME; MULTIGENE FAMILY.
 FT BINDING 40 40 HEME LIGAND (BY SIMILARITY).
 FT BINDING 64 64 HEME LIGAND (BY SIMILARITY).
 FT TRANSMEM 107 127 POTENTIAL.
 FT CONFLICT 10 11 LA -> EF (IN G19853).
 FT CONFLICT 105 105 MISSING (IN G19853).
 SQ SEQUENCE 136 AA; 14979 MW; 6F34BDD9 CRC32;
 Query Match 5.9%; Score 209; DB 1; Length 136;
 Best Local Similarity 41.8%; Pred. No. 3.51e-18;
 Matches 33; Conservative 17; Mismatches 27; Indels 2; Gaps 2;
 Db 1 MGGETKVTLLAE-VSQHNNAKDCWLIVSGKYVDYTKFLDDHPPGGDEVLLSATGKDADF 59
 QY 1 MEGEAKYITAEEDLRNRKSGDLWISIQGVKDCSRWAAEHPPGVPLLSLAGQDVDAF 60
 Db 60 EDVGHSSSARAMLDEYVVG 78
 QY 61 IAY-HPGTAWRHLDPLETG 78
 RESULT 7
 ID CYB5_YEAST STANDARD; PRT; 120 AA.
 AC P40312;
 DT 01-FEB-1995 (REL. 31, CREATED)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)

RESULT	10
--------	----

```

RESULT      11
ID          NIA2_PHAVU    STANDARD;       PRT;        890 AA.
AC          P39866;
DT          01-FEB-1995 (REL. 31, CREATED)
DT          01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT          01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE          NITRATE REDUCTASE 2 (EC 1.6.6.1) (NR-2).
GN          NIA2 OR NR2.
OS          PHASEOLUS VULGARIS (KIDNEY BEAN) (FRENCH BEAN).
OC          EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE; FABACEAE.
CC          [1]
RN          SEQUENCE FROM N.A.
RC          STRAIN=CV. SAXA;
RA          JENSEN P.E., HOFF T., STUMMANN B.M., HENNINGSEN K.W.;
RL          SUBMITTED (AUG-1993) TO EMBL/GENEANK/DDBJ DATA BANKS.
CC -!- FUNCTION: NITRATE REDUCTASE IS A KEY ENZYME INVOLVED IN THE FIRST STEP OF NITRATE ASSIMILATION IN PLANTS, FUNGI AND BACTERIA.
CC -!- CATALYTIC ACTIVITY: NADH + NITRATE = NAD(+) + NITRITE + H(2)O.
CC -!- COFACTOR: REQUIRES FAD, A HEME GROUP (CALLED CYTOCHROME B-557) AND ONE MOLYBDENUM ATOM.
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SIMILARITY: TO EUKARYOTIC MOLYBDOPTERIN OXIDOREDUCTASES IN THE N-TERMINAL DOMAIN.
CC -!- SIMILARITY: CONTAINS A CYTOCHROME B5 FAMILY, HEME-BINDING DOMAIN.
CC -!- SIMILARITY: TO FAD/NAD-BINDING CYTOCHROME REDUCTASES IN THE C-TERMINAL DOMAIN.
DR          EMBL; U01029; G392992; -.
DR          PROSITE; PS00191; CYTOCHROME_B5; 1.
DR          PROSITE; PS00559; MOLYBDOPTERIN_EUK; 1.
KW          OXIDOREDUCTASE; FLAVOPROTEIN; FAD; NAD; HEME; MOLYBDENUM; NITRATE ASSIMILATION; MULTIGENE FAMILY.
FT          METAL     165   165   MOLYBDENUM-PTERIN (POTENTIAL).
FT          METAL     219   219   MOLYBDENUM-PTERIN (POTENTIAL).
FT          DISULFID   404   404   INTERCHAIN (POTENTIAL).
FT          BINDING    548   548   HEME LIGAND (BY SIMILARITY).
FT          BINDING    571   571   HEME LIGAND (BY SIMILARITY).
SQ          SEQUENCE    890 AA; 99995 MW; D952FE9D CRC32;

Query Match           5.7%; Score 202; DB 1; Length 890;
Best Local Similarity 34.2%; Pred.No.6.78e+16;
Matches              25; Conservative 21; Mismatches 27; Indels 1; Gaps

Db  509 MNTASKWF-SVSEVKKHSSPDSAWIIIVGHVYDCRFLKDHPGGTDSILIINAGTDCTEEF 567
QY   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   1 MEGEAKKYITAEDLRHNKSGLDWISOGKVYDCRSWAAEHPGVEPLLSLAGQVDVTDAF 60

Db  568 DAHSKDKAKMKLE 580
QY   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY   61 IAYHPGTAWRHLD 73

RESULT      12
ID          NIA_LYCES    STANDARD;       PRT;        911 AA.
AC          PI7570;
DT          01-AUG-1990 (REL. 15, CREATED)
DT          01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT          01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
DE          NITRATE REDUCTASE (EC 1.6.6.1) (NR).
GN          NIA.
OS          LYCopersicon ESCULENTUM (TOMATO).
OC          EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE; SOLANALES; SOLANACEAE.
CC          [1]
RN          SEQUENCE FROM N.A.
RC          STRAIN=CV. MANAPAL; TISSUE=LEAF;
RA          MEDLINE; 90185211.
RX          DANIEL-VÉDELLE F., DORBE M.F., CABOCHE M., ROUZÉ P.;
RL          GENE 85:371-380(1989).
CC -!- FUNCTION: NITRATE REDUCTASE IS A KEY ENZYME INVOLVED IN THE FIRST STEP OF NITRATE ASSIMILATION IN PLANTS. FUNGI AND BACTERIA.
```

Query Match 5.7%; Score 202; DB 1; Length 918;
Best Local Similarity 37.0%; Pred. No. 6.78e-17;
Matches 27; Conservative 18; Mismatches 27; Indels

Db 539 MNTASNTY-TLSEVKKHNPSQSAIIIVHGVYDCTREFLDHPGGSDSILINAGTDOCTEEF 597
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
QV 1 MEGEAKKYTAEDLRHNKSGLDWISIOGRKVDCSWAAEHGPGGEVPLLSLAGOVDTAF 60

Db 598 DAIHSDKAKKMLE 610

QV 61 IAYHPGTAWRHLD 73

RESULT 14
ID NIA_SPIOL STANDARD; PRT; 926 AA.

DT	01-NOV-1991 (REL. 20, CREATED)
DT	01-NOV-1991 (REL. 20, LAST SEQUENCE UPDATE)
DT	15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE	NITRATE REDUCTASE (EC 1.6.5.1) (NR).

Db 591 DAIHSDAKAKLLE 603

OV 61 IAYHPGTAWRHL 73

RX MEDLINE; 91355858.
RA PROSSER I.M., LAZARUS C.M.;
RL PLANT MOL. BIOL. 15:187-190(1990).

FLANI MOL. BIOL. 13:187-190(1990)
[2]
SEQUENCE FROM N.A.
STRAIN=CV. HOYO;

RX MEDLINE; 97273959.
 RA TAMURA N., TAKAHASHI H., TAKEBA G., SATOI T., NAKAGAWA H.;
 RL BIOCHIM. BIOPHYS. ACTA 1338:151-155(1997).

RL BIOCHIM. BIOPHYS. ACTA 1338:151-155(1997).

RN [1]
SEQUENCE FROM N A

RA SHIRAIISHI N., KUBO Y., TAKEBA K., KIYOTA S., SAKANO K., NAKAGAWA H.;
RL PLANT CELL PHYSIOL. 32:1031-1038(1991).
CC -1- FUNCTION: NITRATE REDUCTASE IS A KEY ENZYME INVOLVED IN THE FIRST
CC STEP OF NITRATE ASSIMILATION IN PLANTS, FUNGI AND BACTERIA.

RN [2]

RP SEQUENCE FROM N.A.
RC TISSUE=SEEDLING:

RX MEDLINE; 92084635.

RL J. BIOL. CHEM. 266:23542-23547(1991).

CC -!- FUNCTION: NITRATE REDUCTASE IS A KEY ENZYME INVOLVED IN THE FIRST STEP OF NITRATE ASSIMILATION IN PLANTS. FUNCT AND DICTERIA

CC -!- CATALYTIC ACTIVITY: NADH + NITRATE = NAD(+) + NITRITE + H(2)O.

CC -!- COFACTOR: EACH SUBUNIT OF THE ENZYME CONTAINS 1 EQUIVALENT OF FAD

[illegible]

CC -!- SUBUNIT: HOMODIMER.

CC -!- INDUCTION: BY NITRATE.
CC -!- SIMILARITY: TO EUKARYOTIC MOLYBDOPTERIN OXIDOREDUCTASES IN THE

CC N-TERMINAL DOMAIN.

CC -!- SIMILARITY: CONTAINS A CYTOCHROME B5 FAMILY, HEME-BINDING DOMAIN.
CC -!- SIMILARITY: TO FAD/NAD-BINDING CYTOCHROME REDUCTASES IN THE

CC C-TERMINAL DOMAIN.

DK EMBL; M33134; G167499; -.
DR PTR: A41667: A41667.

DR HSSP; P17571; 2CND.

DR PROSTIE; PS00191; CYTOCHROME_B5; 1.

Qy 61 IAYHPGTAWRHL D 73

Search completed: Thu Feb 18 11:36:36 1999
Job time : 30 secs.

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Run on: Thu Feb 18 11:36:54 1999; MasPar time 25.99 Seconds
865.947 Million cell updates/sec

Title: >US-08-934-254-27
Description: (1-452) from US08934254.pap
Perfect Score: 3515
Sequence: 1 MEGAKKYITAEDLRHNSK.....LNSAPCPKKGGEAYNTHG 452

Scoring table: PAM 150
Gap 11

Searched: 165420 seqs, 49795644 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: spiremb6

1:sp.archaea 2:sp.bacteria 3:sp.fungi 4:sp.human
5:sp.invertebrate 6:sp.mammal 7:sp.mhc 8:sp.organelle
9:sp.phage 10:sp.plant 11:sp.rodent 12:sp.unclassified
13:sp.vertibrate 14:sp.virus

Statistics: Mean 49.519; Variance 96.434; scale 0.514

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2433	69.2	458	10	CYTCHROME B5 CONTAINI	0.00e+00
2	2133	60.7	448	10	DELTA 6 DESATURASE.	0.00e+00
3	431	12.3	444	4	BC269730.2.	1.11e-58
4	407	11.6	443	5	DELTA6-FATTY-ACID-DESA	6.23e-54
5	405	11.5	473	5	W08D2.4.	1.54e-53
6	317	9.0	454	5	T13F2.1.	1.70e-36
7	220	6.3	875	10	NITRATE REDUCTASE (FRA	1.05e-18
8	207	5.9	134	10	CYTCHROME B5 (FRAGMENT	1.91e-16
9	200	5.7	911	10	NADH NITRATE REDUCTASE	3.04e-15
10	200	5.7	911	10	NADH NITRATE REDUCTASE	3.04e-15
11	194	5.5	113	10	CYTCHROME B5 (FRAGMEN	3.17e-14
12	194	5.5	134	10	PUTATIVE CYTOCHROME B5	3.17e-14
13	193	5.5	427	2	HYPOTHETICAL 48.4 KD P	4.68e-14
14	193	5.5	915	10	NITRATE REDUCTASE APOE	4.20e-14
15	189	5.4	368	2	DELTA 6 DESATURASE.	2.68e-13
16	190	5.4	891	3	NITRATE REDUCTASE (NAD	1.50e-13
17	188	5.3	74	10	NITRATE REDUCTASE HEME	3.24e-13
18	187	5.3	146	4	CYTCHROME B5 (FRAGMEN	4.76e-13
19	185	5.3	497	10	NITRATE REDUCTASE (EC	1.03e-12
20	187	5.3	501	10	NITRATE REDUCTASE (EC	4.76e-13

21	185	5.3	629	10	P92920	NITRATE REDUCTASE (EC	1.03e-12
22	183	5.2	94	10	Q39966	NITRATE REDUCTASE (FRA	2.21e-12
23	182	5.2	95	10	Q39993	NITRATE REDUCTASE (FRA	3.23e-12
24	182	5.2	95	10	Q39976	NITRATE REDUCTASE (FRA	3.23e-12
25	182	5.2	95	10	Q39984	NITRATE REDUCTASE (FRA	3.23e-12
26	183	5.2	132	10	O24651	CYTCHROME B5 (FRAGMEN	2.21e-12
27	183	5.2	132	10	O04354	CYTCHROME B5.	2.21e-12
28	184	5.2	487	10	P92921	NITRATE REDUCTASE (EC	1.51e-12
29	181	5.1	94	10	Q41318	NITRATE REDUCTASE (FRA	3.14e-11
30	176	5.0	91	10	Q38791	NITRATE REDUCTASE (FRA	3.14e-11
31	177	5.0	95	10	Q39971	NITRATE REDUCTASE (FRA	2.15e-11
32	177	5.0	95	10	Q39992	NITRATE REDUCTASE (FRA	2.15e-11
33	177	5.0	95	10	Q41613	NITRATE REDUCTASE (FRA	2.15e-11
34	176	5.0	907	3	Q92237	NITRATE REDUCTASE.	3.14e-11
35	174	5.0	907	3	Q00303	NITRATE REDUCTASE.	6.66e-11
36	171	4.9	98	6	O28726	SOLUBLE CYTOCHROME B5.	2.04e-10
37	167	4.8	171	11	P70116	CYTCHROME B5.	9.02e-10
38	169	4.8	892	3	O13486	NITRATE REDUCTASE.	4.30e-10
39	163	4.6	100	11	O35768	CYTCHROME B5.	3.93e-09
40	163	4.6	121	10	O22704	PUTATIVE CYTOCHROME B5	3.93e-09
41	160	4.6	877	10	O42497	NITRATE REDUCTASE.	1.17e-08
42	157	4.5	864	3	O01696	NITRATE REDUCTASE.	3.48e-08
43	153	4.4	112	5	O17091	CYTCHROME B5.	1.46e-07
44	153	4.4	565	3	O13510	L-MANDELATE DEHYDROGEN	1.46e-07
45	151	4.3	138	5	O93315	C31E10.7.	2.97e-07

ALIGNMENTS

RESULT	ID	PRELIMINARY;	PRT;	458 AA.
AC	Q43469			
DT	01-NOV-1996	(TREMBREL. 01, CREATED)		
DT	01-NOV-1996	(TREMBREL. 01, LAST SEQUENCE UPDATE)		
DT	01-JUN-1998	(TREMBREL. 06, LAST ANNOTATION UPDATE)		
DE	CYTCHROME B5 CONTAINING FUSION PROTEIN.			
OS	HELIANTHUS ANNUUS (COMMON SUNFLOWER).			
OC	EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;			
OC	ASTERALES; COMPOSITAE.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CV. INBRED LINE HA89, SFI63 /91;			
RC	TISSUE=COTYLEDONS OF DEVELOPING SUNFLOWER FRUITS;			
RX	MEDLINE; 96028121.			
RA	SPEERLING P.; SCHMIDT H.; HEINZ E.;			
RL	EUR. J. BIOCHEM. 232:798-805(1995).			
DR	EMBL; X87143; G1040729; -.			
DR	PFAM; PF00173; heme-1.			
SQ	SEQUENCE 458 AA; 52231 MW; FA2AE27B CRC32;			

Query Match	69.2%;	Score 2433;	DB 10;	Length 458;
Best Local Similarity	66.7%;	Pred. No. 0.00e+00;		
Matches	300;	Conservative 92;	Mismatches 54;	Indels 4;
Gaps	4;			
Db	13	ADGKYITKELKKHNPDLWISILGKVVNVTWAKEHGGDAPLNLAGQVTDAPFA 72		
QY	3	GEAKYITAEEDLRHNSKGLWISIQKVVDCSWAAEHGGEVPLLSLAQQVTDAPFA 62		
Db	73	FHPGTANKHDKLFTGVH-LKDYQVSDISDRYKLASEFAKAGMFKKHGHIYSLCFVS 131		
QY	63	YHPGTANKHDKLFTGVH-LKDYQVSDISDRYKLASEFAKAGMFKKHGHIYSLCFVS 122		
Db	132	LLLSACVYGVLYSGFWIHLMSGAILGLAWQIAYLGHGDAGHYOMATRGNNKAGFIF 191		
QY	123	VMMAIYGVGLASEVGHMLCGALLGLWIAQVYGVDSHYQVMPTRGYNRIQLIAG 182		
Db	192	NCITGISIAKWKTHNAHHAACNSLDYDPLQLHPLMLAVSSKLFNSITSVYGRQLTFDP 251		
QY	183	NILTGISIAKWKTHNAHHAACNSLDYDPLQLHPLMLAVSSKLFNSITSVYGRQLTFDP 242		
Db	252	LARFFVSYQHYLYPIMCVARVNLQTLILLISKRIIPDRGLNLTGLIFWTFPLVLS 311		
QY	243	VARELVSYQHYLYPIMCVARVNLQTLILLISKRIIPDRGLNLTGLIFWTFPLVLS 302		

DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE NITRATE REDUCTASE (FRAGMENT)
GN BCNR-A.
OS GLYCINE MAX (SOYBEAN).
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE; FABALES;
OC FABACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 96123229.
RA WU S., LU Q., KRIZ A.L., HARPER J.E.;
RL PLANT MOL. BIOL. 29:491-506(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA CHYNA B., SMARRELLI J.;
RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF022780; G2731816; -.
DR PROSITE; PS00191; CYTOCHROME B5; 1.
DR PROSITE; PS00559; MOLYBDOPTERIN_EUK; 1.
HEME.
NON_TER 1 1
NON_TER 875 875
SQ SEQUENCE 875 AA; 98406 MW; 2CB2F7E7 CRC32;
Query Match 6.3%; Score 220; DB 10; Length 875;
Best Local Similarity 41.1%; Pred. No. 1.05e-18;
Matches 30; Conservative 16; Mismatches 26; Indels 1; Gaps 1;
Db 499 MNTSKTYTMSV-VRRHNNADSAWIIVGHVYDCTFLKDPHGGTDSILINAGTDCTEEF 557
QY 1 MESEAKKYITAEDLRHNSGDLWISIQGVYDCSRWAAEHGPGVEPILLSLAGQDVTDFAF 60
Db 558 EAIHSDKAKOMLE 570
QY 61 IAYHPTAWRHLD 73
RESULT 8
ID 048618 PRELIMINARY; PRT; 134 AA.
AC 048618;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE CYTOCHROME B5 (FRAGMENT).
GN CYTB5-2.
OS OLEA EUROPAEA (COMMON OLIVE).
OC EUKARYOTAE; MITOCHONDRIAL EUKARYOTES; VIRIDIPLANTAE;
CHAROPHYTA/EMBRYOPHYTA GROUP; EMBRYOPHYTA; VASCULAR PLANTS;
SEED PLANTS; MAGNOLIOPHYTA; MAGNOLIOPSIDA; GENTIANANAE;
SCROPHULARIALES; OLEACEAE; OLEA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. KORONEIKI;
RA MARTINKOVSKAYA A.I.;
RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. KORONEIKI;
RA MARTINKOVSKAYA A.I.;
RA MARTINKOVSKAYA A.I., POGHOSYAN Z.P., HARALAMBIDIS K.G.,
RA HATZIOPOULOS P., MURPHY D.J.;
RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AJ001370; E1216720; -.
FT NON_TER 1
SQ SEQUENCE 134 AA; 15316 MW; 4E866893 CRC32;
Query Match 5.9%; Score 207; DB 10; Length 134;
Best Local Similarity 36.7%; Pred. No. 1.91e-16;
Matches 29; Conservative 20; Mismatches 28; Indels 2; Gaps 2;
Db 1 MASDPKIYVY-EVEVHKDTKDWLVINGKYVDVTFPFMDHDPGGDEVLSATGKDATNDF 59
QY 1 MESEAKKYITAEDLRHNSGDLWISIQGVYDCSRWAAEHGPGVEPILLSLAGQDVTDFAF 60

Db 60 EDVGHSDSAREMMDKYIG 78
QY 61 IAY-HPGTAWRHLDPLFTG 78
RESULT 9
ID 004926 PRELIMINARY; PRT; 911 AA.
AC 004926;
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE NADH NITRATE REDUCTASE (EC 1.6.6.3) (NITRATE REDUCTASE (NADPH)).
GN STNR3.
OS SOLANUM TUBEROSUM (POTATO).
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
OC SOLANALES; SOLANACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-DESIREE;
RA HARRIS N., FOSTER J.M., KUMAR A., DAVIES H.V., WRAY J.L.;
RL SUBMITTED (MAR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- CATALYTIC ACTIVITY: NADPH + NITRATE = NADP(+) + NITRITE + H(2)O.
CC -!- COFACTOR: FAD; HEME; MOLYBDENUM.
DR EMBL; U95317; G1946812; -.
DR PROSITE; PS00191; CYTOCHROME B5; 1.
DR PROSITE; PS00559; MOLYBDOPTERIN_EUK; 1.
DR PFAM; PF00173; heme_1.
DR PFAM; PF00174; oxidoreduced_molyb.
DR PFAM; PF00175; oxidoreduced_fad.
KW OXIDOREDUCTASE; HEME.
SQ SEQUENCE 911 AA; 102609 MW; 44AC790F CRC32;
Query Match 5.7%; Score 200; DB 10; Length 911;
Best Local Similarity 39.7%; Pred. No. 3.04e-15;
Matches 29; Conservative 16; Mismatches 27; Indels 1; Gaps 1;
Db 532 MNTASKMYMSV-VRKHNSDSSAWIIVGHYIDASFELKDPHGGVDSILINAGTDCTEEF 590
QY 1 MESEAKKYITAEDLRHNSGDLWISIQGVYDCSRWAAEHGPGVEPILLSLAGQDVTDFAF 60
Db 591 DAHSDKAKKLE 603
QY 61 IAYHPTAWRHLD 73
RESULT 10
ID 024390 PRELIMINARY; PRT; 911 AA.
AC 024390;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE NADH NITRATE REDUCTASE (EC 1.6.6.3).
GN STNR2.
OS SOLANUM TUBEROSUM (POTATO).
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
OC SOLANALES; SOLANACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. DESIREE;
RA HARRIS N., FOSTER J.M., KUMAR A., DAVIES H.V., WRAY J.L.;
RL SUBMITTED (OCT-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; U76701; G1673613; -.
DR PROSITE; PS00191; CYTOCHROME B5; 1.
DR PROSITE; PS00559; MOLYBDOPTERIN_EUK; 1.
DR PFAM; PF00173; heme_1.
DR PFAM; PF00174; oxidoreduced_molyb.
DR PFAM; PF00175; oxidoreduced_fad.
KW OXIDOREDUCTASE; HEME.
SQ SEQUENCE 911 AA; 102596 MW; 6AECEEC9 CRC32;
Query Match 5.7%; Score 200; DB 10; Length 911;
Best Local Similarity 39.7%; Pred. No. 3.04e-15;
Matches 29; Conservative 16; Mismatches 27; Indels 1; Gaps 1;

[illegible]

RESULT	11	PRELIMINARY;	PRT;	113 AA.
ID	Q42342			
AC	Q42342;			
DT	01-NOV-1996	(TREMBREL. 01, CREATED)		
DT	01-NOV-1996	(TREMBREL. 01, LAST SEQUENCE UPDATE)		
DT	01-AUG-1998	(TREMBREL. 07, LAST ANNOTATION UPDATE)		
DE	CTCCHROME B5 (FRAGMENT)			
OS	ARABIDOPSIS THALIANA (MOUSE-EAR CRESS)			
OC	EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONAE;			
OC	CAPPARALES; CRUCIFERAE.			
OC	[1]			
OC	SEQUENCE FROM N.A.			

RC	TISSUE=CELL SUSPENSION CULTURE OF ECOTYPE COLUMBIA;
RD	INSTRUMENT=
RE	COOKE R., LAUDIE M., RAYNAL M., DELSENY M.;
RF	SUBMITTED (MAR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RL	COMMITTEE (F20001; E225588; -;
RR	EMBL: F20001; E225588; -;
DR	DR PFAM: PF00173; heme_1
DR	PFAM: PF00173; heme_1
FT	NON_TER 1
FT	NON_TER 113
FT	NON_TER 113
SQ	SEQUENCE 113 AA; 12686 MW; 66CD71FB CRC32;

Query Match 5.5%; Score 194; DB 10; Length 113;
Best Local Similarity 40.5%; Pred. No. 3.17e-14;
Matches 30; Conservative 16; Mismatches 27; Indels

```

Db      5  RKVLSEFEYSKHNKTKDCWLLISGVKVDVTFPMDDHPGGDEVLLSTGKDANDEDFVGH 64
QY      6  KKITIADERRINKSGDLWISIQGVYDCSRWAAEHPPGGEVPLLVLAGQDVTDAFIAY-H 64

Db      65  SDTARDMDKVFYF 78
QY      65  PGTAWRHLDPLFTG 78

```

RESULT	12
ID	O48845
AC	O48845;
PRT;	134 AA.
DT	01-JUN-1998 (TRENBLREL. 06, CREATED)
DT	01-JUN-1998 (TRENBLREL. 06, LAST SEQUENCE UPDATE)
DT	01-AUG-1998 (TRENBLREL. 07, LAST ANNOTATION UPDATE)
DT	[REDACTED] PUTATIVE CYTOCHROME B5.
GN	F24L7.14.
OS	ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
OC	EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
CC	CAPPARALES; CRUCIFERAE.

[1]
 RN SEQUENCE FROM N.A.
 RP
 RC STRAIN-CV, COLUMBIA;
 RA ROUNSELEY S.D., LIN X., KETCHUM K.A., CROSBY M.L., BRANDON R.C.,
 RA SKES S.M., KAUL S., MASON T.M., KERLAVAGE A.R., ADAMS M.D.,
 RA SOMERVILLE C.R., VENTER J.C.;
 RA SUBMITTED (FEB-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 RL
 DR EMBL; AC003974; G3914701; -
 DR PROSITE; PS00191; CYTOCHROME_B5; 1.
 KW HEME.
 SQ SEQUENCE 134 AA: 15016 MW: 3EAC1E72 CRC32:

Query Match 5.5%; Score 194; DB 10; Length 134;
Best Local Similarity 38.0%; Pred. No. 3.17e-14;
Matches 30; Conservative 16; Mismatches 31; Indels

Db 1 MGDEAKIF-TLSEVSEHNQAHDCWIVNGKYVNVTKFLEDPGGDDVLLSSTGKDATDDF 59

QY	1	MEGEAKKVTIATEDLRRHNKSGLWISIOCKVYDCSRWAAEHPGGVEVPLLSLAGQDVTDAF
Db	60	EDVGHSASAREMEEQYYVG 78 : : :
QY	61	IAY -HPGTAWRHLDPLFTG 78
RESULT	13	
ID	O05874	PRELIMINARY; PRT: 427 AA:

RESULT	13			
ID	O05874	PRELIMINARY;	PRT;	427 AA.
AC	O05874;			
DT	01-JUL-1997	(TREMBLREL. 04, CREATED)		
DT	01-JUL-1997	(TREMBLREL. 04, LAST SEQUENCE UPDATE)		
DT	01-JUL-1997	(TREMBLREL. 04, LAST ANNOTATION UPDATE)		
DE	HYPOTHETICAL 48.4 KD PROTEIN.			
GN	MTCY20B11.04C.			
OS	MYCOBACTERIUM TUBERCULOSIS.			
OC	PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; MYCOBACTERIACEAE.			
[1]				
RN	SEQUENCE FROM N.A.			
RP	STRAIN=H37RV;			
RC	BADCOCK K., CHURCHER C.M.;			
RA	SUBMITTED (MAY-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.			
[2]				
RN	SEQUENCE FROM N.A.			
RP	STRAIN=H37RV;			
RC	COLE S.T., BARRELL B.G., RAJANDREAM M.A.;			
RA	SUBMITTED (MAY-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.			
[3]				
RN	SEQUENCE FROM N.A.			
RP	STRAIN=H37RV;			
RC	MEDLINE; 96181548.			
RX	PHILIPP W.J., POULET S., EIGLMETER K., PASCOPELLA L.,			
RA	BALASUBRAMANIAN V., HEYM B., BERGH S., BLOOM B.R., JACOBS			
RA	COLE S.T.;			
RL	PROC. NATL. ACAD. SCI. U.S.A. 93:3132-3137(1996).			
RL	EMBL; Z95121; E314467; -			
DR	HYPOTHETICAL PROTEIN.			
SW	SEQUENCE 427 AA; 48443 MW; 5D40CA8B CRC32;			
KQ				

	Query Match	5.5%	Score 193;	DB 2;	Length 427;	
	Best Local Similarity	32.9%;	pred. No. 4.68e-14;			
	Matches	26;	Conservative	20;	Mismatches 32;	Indels 1; Gaps 1;
Db	265	FTKDTMTGPKQGWYLRQMLGSANFNAGPALRFMGNTLCHQTEHHLYPDLPSNRRLHEIS	323			
	:	: : :	:	:	:	:
Qy	333	FSGTYVGPCKGDWFEKFKIGIDITCPMWDWFEGGLFQLEHHLFPRLPRGQLRKIA	392			
	:	: : :	:	:	:	:
Db	324	VRREVCDRYDLPYTGTGSF	342			
Ov	393	PLARDICKKHGMPYSRSGF	411			
	:	: : :	:	:	:	:

```

RESULT 14
ID Q43042 PRELIMINARY; PRT; 915 AA.
AC Q43042;
DT 01-NOV-1996 (TREMBREL. 01, CREATED)
DT 01-NOV-1996 (TREMBREL. 01, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBREL. 07, LAST ANNOTATION UPDATE)
DE NITRATE REDUCTASE APOENZYME (EC 1.6.6.1) (NITRATE REDUCTASE (NADH))
DE (ASSIMILATORY NITRATE REDUCTASE).
GN NIA.
OS PETUNIA HYBRIDA (PETUNIA).
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
OC SOLANALES; SOLANACEAE.
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=TLR113; TISSUE=LEAF;
RA SALAMOUNBAT N., BUDANG H.D.;
RL GENE 0:0-0(0).
CC -1- CATALYTIC ACTIVITY: NADH + NITRATE = NAD(+) + NITRITE + H(2)O.
CC -1- COFACTOR: FAD OR FMN; HEME; MOLYBDENUM.
CC EMBL; L11563; G484212; .
DR PROSITE; PS00191; CYTOCHROME B5; 1.

```

```

DR PROSITE: PS00559; MOLYBDOPTERIN_EUK; 1.
DR PFAM: PF00173; heme_1.
DR PFAM: PF00174; oxidored_molyb.
DR PFAM: PF00175; oxidored_fad.
KW OXIDOREDUCTASE; HEME.
SQ SEQUENCE 915 AA; 103037 MW; 17A4F657 CRC32;

Query Match      5.5%; Score 193; DB 10; Length 915;
Best Local Similarity 37.0%; Pred.No. 4.68e-14;
Matches 27; Conservative 18; Mismatches 27; Indels 1; Gaps 1;

Db 537 MNTASKMYSMSP-VKKHNSADSAAVIHGHVYDATRFKLDHPGIDSIILINAGTCTEEF 595
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 1 MEGEAKKVIITAEADLRHKNKSGDLWISIQGVYDCSRWAAEHGGEVPLLSLAGQDVTDAF 60

Db 596 DAIHSDKAKKLE 608
   | | : | : | : |
QY 61 IAYHPGTAWRHLD 73

.T 15
IS Q54795 PRELIMINARY; PRT; 368 AA.
AC Q54795;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-MAY-1997 (TREMBLREL. 03, LAST ANNOTATION UPDATE)
DE DELTA 6 DESATURASE.
GN DESD.
OS SPIRULINA PLATENSIS.
OC PROKARYOTA: GRACILICUTES: OXYPHOTOBACTERIA:
OC CYANOBACTERIA (BLUE-GREEN ALGAE); OSCILLATORIALES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ITALY;
RA MURATA N., DESHNIUM P., TASAKA Y.;
RL (IN) HUANG Y., MILLES D.E. (EDS.);
RL GAMMA-LINOLENIC ACID,
RL METABOLISM AND ITS ROLES IN NUTRITION AND MEDICINE:22-32;
RL AOC PRESS, CHAMPAIGN, ILLINOIS (1996).
DR EMBL; X87094; G809110; -.
SQ SEQUENCE 368 AA; 42083 MW; A9222BFE CRC32;

Query Match      5.4%; Score 189; DB 2; Length 368;
Best Local Similarity 29.7%; Pred.No. 2.20e-13;
Matches 38; Conservative 31; Mismatches 52; Indels 7; Gaps 5;

Db 209 AFKAPGVAVFLI-IPYAVGYSPL-EAVIGASIVYTHGLVACVVFVLAHVIEPAEFLDPD 266
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
284 ALNLGIAVFTWFLFVSCLPNWPFRGFLVISTAVTAIQHVQFTLNH-FSGDTIVGPP 342

Db 267 NLHIDDEWAIQAVKTVDFAPNPNINNYVGLNYQTVVHLPFHICHIHYKPIAIEV 326
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 343 K--GDNWFETKGTIDITCP-PWMDWFFGLQFLEHLPFLPRGQLRKRIAPLARDL 398

Db 327 CEEFGVNY 334
   | | : | : |
QY 399 CKKHGMPY 406

```

Search completed: Thu Feb 18 11:37:56 1999
 Job time : 62 secs.

W P R F (TW)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Mpsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Feb 18 11:33:50 1999; MasPar time 13.14 Seconds
556.518 Million cell updates/sec

ar output not generated.

Title: >US-08-934-254-27
Description: (1-452) from US08934254.pep
Perfect Score: 3515
Sequence: 1 MEGEAKYITAEDLRHNS.....LNSAPCPKILGYEAYNTHG 452

Scoring table: PAM 150
Gap 11

Searched: 131922 seqs, 16180660 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq32

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29

Statistics: Mean 36.009; Variance 155.761; scale 0.231

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description	Pred. No.
1	2134	60.7	448	18	R98455 Borage delta-6-desatu	4.05e-194
2	205	5.8	899	17	R90987 Nitrate reductase.	7.83e-08
3	200	5.7	911	2	R10333 Deduced sequence of t	2.08e-07
4	193	5.5	904	8	R41757 Nitrate reductase Nia	8.15e-07
5	189	5.4	130	23	W22848 Mortierella alpina cy	1.77e-06
6	163	4.6	99	11	R57734 Pre-apo-protein encod	2.52e-04
7	163	4.6	121	11	R57732 Pre-apo-protein encod	2.52e-04
8	163	4.6	156	11	R57733 Bacterial delta-6-des	6.41e-04
9	158	4.5	359	7	R34102 Synchocystis delta-6	1.12e-03
10	135	4.4	437	15	R98456 Fructosyl amino acid	8.96e+00
11	104	3.0	337	25	W24134 IL-4 Stat peptide.	1.24e+01
12	102	2.9	847	15	R88320 Beta-carotene hydroxy	2.79e+01
13	97	2.8	176	3	R13988 CR1 protein	2.79e+01
14	97	2.8	231	1	P92219 Sequence encoded by t	4.49e+01
15	94	2.7	1594	5	P81184 Protein encoded by OR	8.39e+01
16	90	2.6	266	26	W21774 Protein encoded by OR	8.39e+01
17	90	2.6	266	26	W21775 Wild type human epimo	7.19e+01
18	91	2.6	288	13	R66476	

19	91	2.6	288	7	R36549 Human epimorphin.	7.19e+01
20	90	2.6	410	10	R51368 Protein containing Cy	8.39e+01
21	91	2.6	418	11	R60501 Linoleic-acid-desatur	7.19e+01
22	90	2.6	422	15	R88131 Aureobasidin sensitiv	8.39e+01
23	90	2.6	422	15	R88130 Aureobasidin resistan	8.39e+01
24	90	2.6	466	8	R39523 Sequence of gamma-ami	8.39e+01
25	90	2.6	921	5	R28407 CGS-PDE from clone 3C	8.39e+01
26	90	2.6	921	20	W11252 Clone p3CGS-5 cyclic	8.39e+01
27	90	2.6	921	13	R69727 Cyclic-GMP stimulated	8.39e+01
28	90	2.6	921	22	W18048 Cyclic-GMP stimulated	8.39e+01
29	90	2.6	942	13	R69728 Cyclic GMP stimulated	8.39e+01
30	90	2.6	942	20	W11239 Cyclic GMP stimulated	8.39e+01
31	90	2.6	942	22	W18049 CGS-PDE amino acid se	8.39e+01
32	90	2.6	942	5	R28408 CGS PDE from bovine b	8.39e+01
33	91	2.6	1026	27	W32059 Dogfish shark kidney	7.19e+01
34	88	2.5	109	3	P61525 Sequence encoded by s	1.14e+02
35	88	2.5	114	22	W20571 H. pylori secreted or	1.14e+02
36	88	2.5	125	9	R47656 Interferon induced 1-	1.14e+02
37	88	2.5	352	25	W16646 Truncated HSV specifi	1.14e+02
38	88	2.5	360	25	W16645 Truncated HSV specifi	1.14e+02
39	88	2.5	362	25	W16642 Truncated HSV specifi	1.14e+02
40	88	2.5	375	25	W16644 Truncated HSV specifi	1.14e+02
41	88	2.5	379	25	W16641 Truncated HSV specifi	1.14e+02
42	88	2.5	392	25	W16640 Truncated HSV specifi	1.14e+02
43	88	2.5	412	25	W16639 Truncated HSV specifi	1.14e+02
44	88	2.5	438	25	W16638 Truncated HSV specifi	1.14e+02
45	89	2.5	491	2	P70464 Sequence of gpJ encod	9.80e+01

ALIGNMENTS

RESULT 1
ID R98455 standard; Protein; 448 AA.
AC R98455;
DE 15-SEP-1996 (first entry)
DE Borage delta-6-desaturase.
KW Delta-6-desaturase; gamma-linolenic acid; transgenic plant; borage;
KW polyunsaturated fatty acid; octadecatetraenoic acid;
KW Chilling resistance; oilseed.
OS Borage officinalis.
FH Key Location/Qualifiers
FT region 156..163 /label= Lipid_box
FT region 196..200 /label= Metal_box-1
FT region 372..377 /label= Metal_box-2
FT WO9621022-A2.
PN 11-JUL-1996.
PD 28-DEC-1995; IB1167.
PF 30-DEC-1994; US-366779.
PR (RHON) RHONE POULENC AGROCHIMIE.
PI Freyssinet GL, Nuccio M, Nunberg AN, Reddy AS, Thomas TL;
DR WPI; 96-333997/33.
DR N-PSDB; T30395.
PT Transgenic plants comprising the borage delta-6-desaturase gene
PT show increased production of gamma linolenic acid and having
PT increased resistance to chilling
PS Claim 3; Page 52-53; 73pp; English.
CC Borage delta-6-desaturase (R98455) catalyses the conversion of
CC linoleic acid to gamma-linolenic acid (GLA). Its sequence was
CC deduced from that of the delta-6-desaturase gene (T30395) isolated
CC from a borage membrane-bound polysomal library. The sequence is
CC distinct from that of Synchocystis delta-6-desaturase (R98456).
CC Expression of the desaturase in transgenic plants, esp. sunflower,
CC soybean, maize, tobacco, peanut, carrot or oilseed rape, results in
CC increased GLA prodn. Alteration of the plant membrane lipids as a
CC result of expression of the desaturase may also result in increased
CC resistance to chilling.
SQ Sequence 448 AA;
Query Match 60.7%; Score 2134; DB 18; Length 448;
Best Local Similarity 57.8%; Pred. No. 4.05e-194;

Matches 262; Conservative 105; Mismatches 80; Indels 6; Gaps 6;

Db 1 maaqikkyitsdelknhdpgdglwisiqkaydsvdkdhpqggsfpklslagqevtdaf 60
QY 1 MGEAKKIYITAEEDLRHNSGDLWISIQGVYDCSRWAAEHFPGGEVPLLSLAGQDVTDAF 60

Db 61 vafhaestwknldkftgyy-lkdyvsevskdyrklvfefskmglykdkgh-imfatlc 118
QY 61 IAYHPTAWRHLDPLFTGYYIKDFVSEISKDYRRLNEMSRSGIFERKKGHHIMW-TFV 119

Db 119 fiamfamvvyvfcogvhlvlfsgclmgflwsgwighdaghymvvsdrlnkfmgj 178
QY 120 GVAVMAAIVGVGLASESVGVHMLCAGLLGLLWIAQAAVYGHDSGHYQVMPTRGNIITQL 179

Db 179 faancslgisgwknwnhahiaensleydpdlqyplfvsskffsgltsfhfyekrit 238
QY 180 IAGNLTGISAIAWKWTHNAHLACNSLDYDPDLQHIPVFAVSTRLFNSITSVFYGRVLK 239

Db 239 fgsrlrfvysghwtfyplmcaarlmyvgsimlittknvsvyragellcglvsiwyp 298
QY 240 FDEVARFLVSYQHWTYIPVWIFGRVNLFIQTLELLLRDVPDRALNMGIAVFWTWFFPL 299

Db 299 lvsclpnwerimfviaslsvtgmqgvqfslnhfssvvygkpgnnwfeqtdgtldis 358
QY 300 FVSCLPNWERGFLVLSFAVTAIOHQVTLNHFSGDTYVGPCKDNWFEKTKGTIDIT 359

Db 359 cppwmdfhgglqfiehlfkmpcrnlrkispyvieleckhnlpy-nyasfskanent 417
QY 360 CPPWMDWFEGGLQFLEHLLFPLRGQLRKIAPLARDLCKKHGMPYRSGFWDANVRT 419

Db 418 lrtltntalqarditk-plpknlvw-ealhtg 448
QY 420 IFTLRDAVAQARDLNSAPCKPLGLGYEAYNTG 452

RESULT 2

ID R90987 standard; Protein; 899 AA.
AC R90987;
DT 01-AUG-1996 (first entry)
DE Nitrate reductase.
KW Nitrate reductase; populus nigra; absorption; nitrogen oxide; pollutant.
OS Populus nigra.
PN J08023978-A.
PD 30-JAN-1996.
PF 14-JUL-1994; JP-162197.
PR (TOY) TOYOTA CHUO KENYUSHO KK.
PA (TOYT) TOYOTA JTDOSHA KK.
WPI: 96-133419/14.
DR N-PSDB; T12999.

PT Populus sp. nitrate reductase gene - useful for generating
PT transgenic plants with enhanced nitrogen oxide absorbing activity
PS Claim 1; Page 4-8; 9pp; Japanese.
CC This sequence represents the nitrate reductase sequence obtained from
CC Populus nigra L. The gene encoding this sequence can be used to generate
CC plants, especially trees, having enhanced activity for absorbing nitrogen
CC oxide pollutants.
SQ Sequence 899 AA;

Query Match 5.8%; Score 205; DB 17; Length 899;
Best Local Similarity 38.4%; Pred. No. 7.83e-08;
Matches 28; Conservative 17; Mismatches 27; Indels 1; Gaps 1;

Db 527 mtsskktfsmas-vkxhsadsaihvghvdyctfklkdhpggtdslinagtcteeef 585
QY 1 MGEAKKIYITAEEDLRHNSGDLWISIQGVYDCSRWAAEHFPGGEVPLLSLAGQDVTDAF 60

Db 586 daihsdskakmie 598
QY 61 IAYHPTAWRHLD 73

RESULT 3

ID R10333 standard; Protein; 911 AA.
DT 08-APR-1991 (first entry)
DE Deduced sequence of tomato nitrate reductase.
KW tomato nitrate reductase; nitrogen assimilation.
OS Lycopersicon esculentum.
PN EP-409730-A.
PD 23-JAN-1991.
PF 18-JUL-1990; 402077.
PR 19-JUL-1989; FR-009707.
PI (INRG) INST NAT RECH AGRON.
PA Daniel-Vedele F, Caboche M;
WPI: 91-024287/04.
DR Q-PSDB; Q10280.

PT New DNA encoding tomato nitrate reductase - and related cloning
PT and expression vectors, used to improve nitrogen assimilation in
PT plants

PS Claim 2; Fig 1: 27pp; French.
CC An EcoRI digest of tomato DNA was screened with labelled tobacco
CC nitrate reductase cDNA under low stringency conditions. A 6.5kb
CC fragment contained part of the tomato nitrate reductase gene and
CC was used, under high stringency, to screen a Hind III library. A
CC 7kb fragment was isolated, having 1.5kb in common with the EcoRI
CC fragment. Restriction fragments were subcloned and sequenced. The
CC deduced protein sequence is encoded by four exons. Comparison with
CC known nitrate reductases from other plants, e.g. Nicotiana,
CC Solanum, Trifolium pratense, Cucumis sativus, etc., shows conserved
CC regions, these have structural and/or functional importance and
CC tend to be separated by variable regions, often hydrophilic, which
CC are possibly specific surface epitopes of the enzyme.
SQ Sequence 911 AA;

Query Match 5.7%; Score 200; DB 2; Length 911;
Best Local Similarity 39.7%; Pred. No. 2.08e-07;
Matches 29; Conservative 16; Mismatches 27; Indels 1; Gaps 1;

Db 532 mntaskmymse-vrkhnssdsawilvghydhadsfklkdhpggvdslinagtcteeef 590
QY 1 MGEAKKIYITAEEDLRHNSGDLWISIQGVYDCSRWAAEHFPGGEVPLLSLAGQDVTDAF 60

Db 591 daihsdskakmie 603
QY 61 IAYHPTAWRHLD 73

RESULT 4

ID R41757 standard; Protein; 904 AA.
AC R41757;
DT 21-MAR-1994 (first entry)
DE Nitrate reductase Nia2 derivative.
KW Nitrate reductase; germination; flowering; ripening; development;
KW growth stimulation; Agrobacterium tumefaciens; nitrate.
OS Nicotiana sp.
PN W09318154-A.
PD 16-SEP-1993.
PF 05-MAR-1993; F00222.
PR 05-MAR-1992; FR-002658.
PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
PI Caboche M, Chupau Y, Dorlhac F, Morot-gaudry J;
WPI: 93-303468/38.
DR P-PSDB; Q48468.

PT Inducing over-expression of nitrate reductase in plants - esp. by
PT incorporation of foreign gene, for stimulating early development
PT and reducing nitrate accumulation
PS Disclosure; Figure 3; 37pp; English.
CC Overexpression of nitrate reductase (NR) can stimulate the early
CC development of plants, shortening the duration of the vegetative
CC phase and causing earlier germination, flowering and ripening by
CC about two weeks. Overexpression of NR can also cause the level of
CC nitrate stored in a plant to be reduced, reducing risks to health
CC and also possibly improving organoleptic qualities. The NR gene was
CC introduced into plants by transforming a strain of Agrobacterium

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

RESULT 8

ID R57733 standard; Protein; 156 AA.
 AC R57733; 1995 (first entry)
 DE Pre-apo-protein encoded by plasmid pSEC-cyt/c.
 KW Plasmid pSEC-cyt/c; vector; halo-protein; pre-apo-protein;
 KW protein synthesis; Escherichia coli; cytoplasm cytochrome.
 OS Synthetic.
 PN WO941191-A.
 PD 04-AUG-1994.
 PF 27-JAN-1994; G00161.
 PR 27-JAN-1993; GB-001553.
 PA (UYWA-) UNIV COLLEGE WALES.
 PI Kaderbhai MA.
 DR WPI; 94-264104/32.
 DR N-PSDB; Q67221.
 PT Genetic precursor unit causing periplasmic translocation of
 PT pre-apo-protein - for processing the halo-protein, useful in
 PT protein synthesis, assay of signal peptidase and identification
 PT of its inhibitors.
 PT Claim 10; Fig.4; 32pp; English.
 CC This sequence encodes the pre-form of an apo-protein. The pre-apo-
 CC protein is translocated from a cytoplasmic cell region to a
 CC periplasmic region where formation of process apo-protein and
 CC conversion to halo-protein occur. The protein is expressed in the
 CC cytoplasm of E. coli, especially TB-1 and N4830-1. The apo-protein
 CC is a cytoplasmic cytochrome, especially having a soluble core domain
 CC of cytochrome-b5 of liver endoplasmic reticulum.
 SQ Sequence 156 AA;
 Query Match 4.6%; Score 163; DB 11; Length 156;
 Best Local Similarity 37.1%; Pred.No. 2.52e-04;
 Matches 26; Conservative 13; Mismatches 31; Indels 0; Gaps 0;
 Db 32 kytleeiqkhdskstwillhkvdytkfleehpggeevlregagsgdatenfedvghs 91
 QY 7 KYTAEELRRNKSGDLWISQGVYDCSRWAEHPGGEVPLLSLAGQDVTDATIAHPG 56
 Db 92 tdarelskty 101
 QY 67 TAWRHLDPLF 76
 RESULT 9
 ID R34102 standard; Protein; 359 AA.
 AC R34102;
 DT 04-AUG-1993 (first entry)
 DE Bacterial delta-6-desaturase.
 KW Gamma-linolenic acid; GLA; linoleic acid; diet; chilling tolerance;
 OS plant.
 OS Synchocystis (PCC 6803, ATCC 27184).
 PN WO9306712-A.
 PD 15-APR-1993.
 PF 13-OCT-1992; U08746.
 PR 10-OCT-1991; US-774475.
 PR 08-JAN-1992; US-817919.
 PA (RHON) RHONE POULENC AGROCHIMIE.
 PI Freyssinet G, Nuccio M, Reddy AS, Thomas T;
 DR WPI; 93-134023/16.
 DR N-PSDB; Q40057.
 PT Nucleic acid encoding bacterial delta-6-desaturase and transgenic
 PT plants - for inducing prodn. of gamma-linolenic acid and
 PT improving chilling resistance of plants
 PS Disclosure: Page 30-33; 45pp; English.
 CC Providing the gene (I) encoding delta-6-desaturase, allows prodn.
 CC of transgenic organisms which contain functional delta-6-desaturase
 CC and which produce gamma-linolenic acid (GLA). In addition to allowing
 CC prodn. of large amts. of GLA, a new dietary source of GLA is provided.
 CC Chilling tolerance may be introduced in plants by introduction of (I)
 CC into a plant cell and regeneration of a plant with improved chilling
 CC resistance from the transformed plant cell.
 SQ Sequence 359 AA;
 Query Match 4.5%; Score 158; DB 7; Length 359;
 Best Local Similarity 37.1%; Pred.No. 2.52e-04;
 Matches 26; Conservative 13; Mismatches 31; Indels 0; Gaps 0;
 Db 32 kytleeiqkhdskstwillhkvdytkfleehpggeevlregagsgdatenfedvghs 91
 QY 7 KYTAEELRRNKSGDLWISQGVYDCSRWAEHPGGEVPLLSLAGQDVTDATIAHPG 56
 Db 92 tdarelskty 101
 QY 67 TAWRHLDPLF 76

Best Local Similarity 31.6%; Pred.No. 6.41e-04;
 Matches 25; Conservative 16; Mismatches 35; Indels 3; Gaps 3;
 Db 273 dewaicqirttanfatnmpfw-nwfcglnhqvthlfpnchihypglenikdvceqef 331
 QY 345 DNWFEKQTGTIDI-TCPP-WMDWFFGGLQFLEHHLFPRPQRLRKIAPLARDLCKKH 402
 Db 332 gveykvptfkaaiastyr 350
 QY 403 GMPYRSFGFWDANVTIR 421
 RESULT 10
 ID R98456 standard; Protein; 359 AA.
 AC R98456;
 DT 15-SEP-1996 (first entry)
 DE Synchocystis delta-6-desaturase.
 KW Delta-6-desaturase; gamma-linolenic acid; transgenic plant; borage;
 KW polyunsaturated fatty acid; octadecatetraenoic acid;
 KW chilling resistance; oilseed.
 OS Synchocystis sp. strain PCC 6803 (ATCC 27184).
 FH Key Location/Qualifiers
 FT region 85..92
 FT /label= Lipid_box
 FT region 123..128
 FT /label= Metal_box-1
 FT region 302..307
 FT /label= Metal_box-2
 PN WO9621022-A2.
 PD 11-JUL-1996.
 PF 28-DEC-1995; IB1167.
 PR 30-DEC-1994; US-366779.
 PA (RHON) RHONE POULENC AGROCHIMIE.
 PI Freyssinet G, Nuccio M, Nunberg AN, Reddy AS, Thomas T;
 DR WPI; 96-333997/33.
 DR N-PSDB; T30396.
 PT Transgenic plants comprising the borage delta-6-desaturase gene -
 PT show increased production of gamma linolenic acid and having
 PT increased resistance to chilling
 PS Disclosure: Page 48-49; 75pp; English.
 CC Synchocystis sp. PCC 6803 delta-6-desaturase (R98456) catalyses
 CC the conversion of linoleic acid to gamma-linolenic acid (GLA). Its
 CC sequence was deduced from that of the delta-6-desaturase gene
 CC (T30396) isolated from a Synchocystis cosmid genomic library.
 CC The sequence is distinct from that of borage delta-6-desaturase
 CC (R98455). Expression of the desaturase in bacteria, fungi,
 CC animals or plants results in increased GLA prodn. Alteration of
 CC the transgenic plant membrane lipids as a result of expression of
 CC the desaturase may result in increased resistance to chilling.
 SQ Sequence 359 AA;
 Query Match 4.4%; Score 155; DB 18; Length 359;
 Best Local Similarity 32.8%; Pred.No. 1.12e-03;
 Matches 22; Conservative 16; Mismatches 26; Indels 3; Gaps 3;
 Db 273 dewaicqirttanfatnmpfw-nwfcglnhqvthlfpnchihypglenikdvceqef 331
 QY 345 DNWFEKQTGTIDI-TCPP-WMDWFFGGLQFLEHHLFPRPQRLRKIAPLARDLCKKH 402
 Db 332 gveykvty 338
 QY 403 GMPYRSF 409
 RESULT 11
 ID W24134 standard; Protein; 437 AA.
 AC W24134;
 DT 16-FEB-1998 (first entry)
 DE Fructosyl amino acid oxidase.
 KW Fructosyl amino acid oxidase; FAOD-P; enzyme; amadori compound;
 KW alpha-ketoaldehyde amine derivative; amadori detection.
 OS Penicillium janthinellum.
 PN WO9721818-A1.

```

Query Match          2.9%; Score 102; DB 15; Length 847;
Best Local Similarity 21.1%; Pred.No. 1.24e+01;
Matches 24; Conservative 36; Mismatches 44; Indels 10; Gaps 9;

Db 433 sendryp-fvaarvvpwkmcelnlhkfnaevtnrgllp-ehflflaqlkfndnslsme 490
      :::::| | | : : : : : | : : : | : : : | : : | : : | : : | : :
QY 211 PDLQHPVFAVSNRL-FNSITSFYGRVLKDFEVARFLVSQHWTYTPVMIFGRVNLFIQ 269
      :| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 491 afghrsvsqfkekl-llgrgtfqwgfdgvidlktkrlsrswsdrlligfls 543
      :| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 270 TFLLLTRRDVPRALNLGIA-VFTWTFP-L-FVS-CLPN-WPERFGFVLIS 317
      :| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
ID ID R1398 standard; Protein; 176 AA.
AC R1398;
DT 26-NOV-1991 (first entry)
DE Beta-carotene hydroxylase - variant.
KW GGPP; carotenoid; phytoene; zeaxanthin; lycopene; ss.
OS Erwinia herbicola EHO-10 (E. vulneris - ATCC 39368).
PN WO91113078-A.
PD 05-SEP-1991.
PF 04-AUG-1991; U01458.
PR 02-MAR-1990; WO-487613.
PR 18-MAY-1990; US-525551.
PR 03-AUG-1990; US-562674.
PA (STAD ) AMOCO CORP.
PI Ausich RL, Brinkhaus FL, Mukharji I, Proffitt JH, Yarger JG;
PI Yen HC;
DR WPI; 91-281410/38.
DR N-PSDB; Q13725.
PT Biosynthesis of carotenoid(s) in genetically engineered hosts -
PT using DNA encoding enzymes from Erwinia herbicola
PS disclosure; Fig 2(1-3); 313pp; English.
CC There are a total of six relevant genes in a 7900 bp region that
CC cause E. coli cells to produce GGPP and the carotenoids phytoene
CC through zeaxanthin diglucoside, which is the final prod. identified
CC in the carotenoid pathway contd. in plasmid pARC376 (contg. a ca.
CC 13 kb chromosomal DNA fragment isolated by Perry et al., J. Bacteriol.,
CC 168:507 (1986)). The genes for geranylgeranyl pyrophosphate (GGPP)
CC synthase, phytoene synthase, phytoene dehydrogenase-4H, lycopene
CC cyclase, beta-carotene hydroxylase, and zeaxanthin glycosylase are
CC represented in Q13716, Q13718, Q13719, Q13722, Q13724 and Q13726
CC respectively.
CC The native sequence (Q13724) was genetically engineered.
CC At the 5' end of the gene, codons encoding the native second and
CC third amino acid have been changed from Leu-val to Val-Leu.
CC Recombinant expression plasmids can be used to produce large amts.
CC of the enzymes and hence large amts. of the carotenoids which they
CC synthesise.
SQ Sequence 176 AA;

Query Match          2.8%; Score 97; DB 3; Length 176;
Best Local Similarity 28.8%; Pred.No. 2.79e+01;
Matches 23; Conservative 22; Mismatches 30; Indels 5; Gaps 5;

Db 38 hhtprkvgfelndlfav-fvagvaialiavgtagvplqwigcmtyvglllyfl-vhdgl 95
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 99 NEMSRGIFPKKGHHIMWTIFGVGVMMAAI-VYGVLAESVGVHMLCGALLGLLTQAAAY 157
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 96 v-hqrpfhwpirgylkrl 114
      | : : : : | : : : :
QY 158 VGHDSGHYQVMPTRGY-NRI 176
      | : : : : | : : : :

RESULT 14
ID ID P92219 standard; protein; 2317 AA.
AC P92219;
DT 22-FEB-1990 (first entry)
DE CR1 protein
KW Complement; cofactor.
OS Homo sapiens (human).
FH Key Location/Qualifiers

```


peptide 10..50
/label= signal_peptide

WO8909220-A.
05-OCT-1989.
31-MAR-1989. PF
01-APR-1988. US-176532.
(CELL) T Cell Sciences, Inc; (UJJO) The Johns Hopkins University;
PA (BRIG*) The Brigham and Women's Hospital.
FI Fearon DT, Klickstein LB, Wong W, Carson G, Makrides SC;
PI WPI: 89-309498/42.
DR N-PSDB: N91477.
PT New nucleic acid sequences encoding new CR1 protein - and its fragment,
PT for diagnosis and control of complement-related immune defects,
PT inflammation, myocardial infarct, etc
PS Claim 1: fig. 1: 191pp; English.
CC This is full-length CR1 protein, and shortened forms are new, lacking
CC the transmembrane region. The proteins and fragments bind C3b and/or
CC C4b, have cofactor activity and inhibit C3 and C5 convertase activity.
CC In the sequence, x-untranslated region. This has 7 short consensus
CC repeats (CR) of 4 long direct homologous repeats (LHR) (see fig. 10 in
CC patent). Deletion mutants, eg LHR-A for binding of C4b, and LHR-C
CC for C3b binding, be made, and LHR-B and -C can detect CR2 sequences.
CC They are useful in diagnosing and treating immune disorders, and prevent
CC perfusion injury.
SQ Sequence 2317 AA:

	Query Match	2.8%	Score 97	DB 1	Length 2317
	Best Local Similarity	27.9%	Pred. No. 2.79e+01		
	Matches	24	Conservative	19	Mismatches 38
				Indels	5
				Gaps	5
Db	2192	nyatyxxpllifxifcl-ssfyvff-yirtfpplvcfillfyr-epysfkqf-rvkym	2247		
	:	:	:	:	:
	:	:	:	:	:
Qy	208	DYDPDLQHPVFAVSTRLENSITSVFYGRVLKDEVARFLVSQHWHTYYPVMIFGRVNL	267		
Db	2248	lyqfllslgrkinallerhemigrv	2273		
	:	:	:	:	:
Qy	268	IQFLLLLLTR-DVPPRALNMGIAV	292		
	:	:	:	:	:

RESULT 15
ID P81184 standard; Protein; 1594 AA.
AC P81184;
DE 29-OCT-1992 (first entry)
DE Sequence encoded by the 2nd reading frame of the peplomeric gene
DE of FIVP strain 79-1146.
KW vaccine; peplomeric protein; E2 gene.
KW Feline infectious peritonitis virus.
CC EF-264979-A.
CC 27-APR-1988.
PR 01-SEP-1987; 201657.
PR 05-SEP-1986; NL-002244.
PR (DUIN) DUFHAR INT RES BV.
PI De Groot RJ, Spaan WJM, Van Der Zeijst BAM;
PI WPI; 88-114147/17.
PR N-PSDB; N81533.
PT Gene for feline infectious peritonitis virus - and gene prod.
PT useful as antigenic protein for vaccine
PS Disclosure; Fig 1; 13pp; English.
CC CDNA was prepd. from FIVP strain 79-1146. N81533 gives the sequence
CC of the peplomeric gene in three reading frames. The top reading
CC frame is an open reading frame of 4356 nucleotides and has a coding
CC capacity for a precursor polypeptide having a mol. wt. of 160,470
CC (1452 AAs). The beginning and the end of the E2 gene are indicated
CC in the FT of N81533. The first 18 N-terminal AAs have a strong
CC hydrophobic character and presumably comprise a cleavable signal
CC peptide. The extreme carboxy-terminal part comprises a region of 20
CC hydrophobic AAs, which presumably serves as a transmembrane anchor.
CC The FIVP peplomeric protein has 35 potential glycosylation sites,
CC of which 22 are in the N-terminal part (pos. 1-790) which corresponds
CC to the S-part of the IBV E2 (see P81183). N.B. IBV = infectious
CC bronchitis virus. "X" in the AA sequence denotes the translation
CC of a stop codon.
SQ Sequence 1594 AA;

	Query Match	2.7%	Score 94;	DB 5;	Length 1594;
	Best Local Similarity	26.4%	Pred. No. 4.49e+01;		
	Matches	24;	Conservative	36;	Mismatches 7; Indels 7; Gaps 7;
Db	817 vl-lldeltvrylvayit-hhyqvicaalkmlvmvsvfilxrhvmxahkrll-lm-vpxlel	872			
	: : : : :	: : : : :	: : : : :		
Qy	237 VLAFDEVA-REFLSYQHWYTPVMTGRVNLFIQTLELLILTRRDVPDRALNMGIAVFWT	295			
Db	873 xlpitvncxvxhigqrhlifittlyitqvr	903			
	: : : : : : :	: : : : : : :			
Qy	296 WFFLEVSC-LPNWPER-FGFVLISFAVTAIQ	324			

Search completed: Thu Feb 18 11:34:58 1999
Job time : 68 secs.

WORLD (TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Feb 18 11:38:16 1999; MasPar time 6.00 Seconds
Jar output not generated. 712.828 Million cell updates/sec

Title: >US-08-934-254-27
Description: (1-452) from US08934254.pep
Perfect Score: 3515
Sequence: 1 MEGEAKYITAEDLRHNS.....LNSAPCPKLGYEAYNTHG 452

Scoring table: PAM 150
Gap 11

Searched: 100342 seqs, 9469514 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-issued
1:5A_COMB 2:5B_COMB 3:PCT9_COMB 4:backfiles1

Statistics: Mean 33.831; Variance 160.640; scale 0.211

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2121	60.3	448	1	US-08-366- Sequence 5, Applicatio	1.93e-176
2	2121	60.3	448	2	US-08-789- Sequence 5, Applicatio	1.93e-176
3	187	5.3	155	2	US-08-801- Sequence 1, Applicatio	6.01e-06
4	171	4.9	134	2	US-08-801- Sequence 4, Applicatio	1.03e-04
5	170	4.8	142	2	US-08-801- Sequence 3, Applicatio	1.22e-04
6	155	4.4	359	1	US-08-366- Sequence 2, Applicatio	1.67e-03
7	155	4.4	359	1	US-08-789- Sequence 2, Applicatio	1.67e-03
8	155	4.4	359	1	US-08-307- Sequence 2, Applicatio	1.67e-03
9	155	4.4	359	1	US-08-478- Sequence 2, Applicatio	1.67e-03
10	155	4.4	359	1	US-08-473- Sequence 2, Applicatio	1.67e-03
11	137	3.9	104	2	US-08-801- Sequence 5, Applicatio	3.57e-02
12	102	2.9	847	1	US-08-781- Sequence 2, Applicatio	1.01e+01
13	102	2.9	847	1	US-08-276- Sequence 2, Applicatio	1.01e+01
14	97	2.8	175	1	US-08-624- Sequence 6, Applicatio	2.17e+01
15	97	2.8	176	1	US-08-096- Sequence 16, Applicatio	2.17e+01
16	91	2.6	288	2	US-08-628- Sequence 3, Applicatio	5.33e+01
17	91	2.6	288	1	US-08-690- Sequence 3, Applicatio	5.33e+01
18	91	2.6	418	3	PCT-US94-0 Sequence 72, Applicatio	5.33e+01
19	93	2.6	450	2	US-08-818- Sequence 6, Applicatio	3.96e+01
20	91	2.6	473	1	US-08-439- Sequence 4, Applicatio	5.33e+01
21	91	2.6	473	1	US-08-440- Sequence 3, Applicatio	5.33e+01
22	93	2.6	553	1	US-08-565- Sequence 6, Applicatio	3.96e+01
23	90	2.6	921	3	PCT-US92-0 Sequence 39, Applicati	6.17e+01

24	90	2.6	921	2	US-08-455- Sequence 39, Applicati	6.17e+01
25	90	2.6	921	1	US-08-297- Sequence 39, Applicati	6.17e+01
26	90	2.6	921	2	US-08-455- Sequence 39, Applicati	6.17e+01
27	90	2.6	921	2	US-08-479- Sequence 39, Applicati	6.17e+01
28	90	2.6	921	1	US-07-872- Sequence 39, Applicati	6.17e+01
29	90	2.6	921	1	US-08-297- Sequence 39, Applicati	6.17e+01
30	90	2.6	942	1	US-07-872- Sequence 43, Applicati	6.17e+01
31	90	2.6	942	1	US-08-297- Sequence 43, Applicati	6.17e+01
32	90	2.6	942	3	PCT-US92-0 Sequence 43, Applicati	6.17e+01
33	90	2.6	942	2	US-08-455- Sequence 43, Applicati	6.17e+01
34	90	2.6	942	1	US-08-297- Sequence 43, Applicati	6.17e+01
35	90	2.6	942	2	US-08-479- Sequence 43, Applicati	6.17e+01
36	90	2.6	942	2	US-08-455- Sequence 43, Applicati	6.17e+01
37	88	2.5	125	3	PCT-US93-0 Sequence 17, Applicati	8.27e+01
38	88	2.5	266	2	US-08-773- Sequence 4, Applicatio	8.27e+01
39	89	2.5	383	3	PCT-US94-0 Sequence 2, Applicatio	7.15e+01
40	89	2.5	516	1	US-08-356- Sequence 4, Applicatio	7.15e+01
41	87	2.5	1078	1	US-08-485- Sequence 7, Applicatio	9.57e+01
42	87	2.5	1078	2	US-08-484- Sequence 7, Applicatio	9.57e+01
43	87	2.5	1078	2	US-08-480- Sequence 7, Applicatio	9.57e+01
44	87	2.5	1085	2	US-08-480- Sequence 5, Applicatio	9.57e+01
45	87	2.5	1088	1	US-08-485- Sequence 6, Applicatio	9.57e+01

ALIGNMENTS

RESULT 1
ID US-08-366-779-5 STANDARD; PRT; 448 AA.
XX
AC xxxxxx
XX
DT
XX
DE Sequence 5, Application US/08366779
XX
XX Sequence 5, Application US/08366779
CC Patent No. 5614393
CC GENERAL INFORMATION:
CC APPLICANT: Thomas, Terry L.
CC APPLICANT: Reddy, Avutu S.
CC APPLICANT: Nuccio, Michael
CC APPLICANT: Freysinet, Georges L.
CC APPLICANT: Nunberg, Andrew N.
CC TITLE OF INVENTION: PRODUCTION OF GAMMA LINOLENIC ACID BY A
CC TITLE OF INVENTION: DELTA 6-DESATURASE
CC NUMBER OF SEQUENCES: 25
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Scully, Scott, Murphy & Presser
CC STREET: 400 Garden City Plaza
CC CITY: Garden City
CC STATE: New York
CC COUNTRY: United States
CC ZIP: 11530
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/366,779
CC FILING DATE: 30-DEC-1994
CC CLASSIFICATION: 800
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Presser, Leopold
CC REGISTRATION NUMBER: 19,827
CC REFERENCE/DOCKET NUMBER: 8383ZYXW
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (516) 742-4343
CC TELEFAX: (516) 742-4366
CC TELEX: 230 901 SANS UR
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 448 amino acids

```

CC      TYPE: amino acid
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: DNA (genomic)
SQ      SEQUENCE 448 AA: 51599 MW: 1110962 CN:

Query Match
Best Local Similarity 60.3%; Score 2121; DB 1; Length 448;
Matches 261; Conservative 105; Mismatches 81; Indels 6; Gaps 6;

Db 1 MAAQIKKTYITDELKNDKPGDLWISIOGKAYDVSDWKDHPGGSPFLKSLAGQEVTDFAF 60
QY 1 MEGEAKKITYITADLRHNKSGDLWISIOGKAYDVSDWKDHPGGSPFLKSLAGQEVTDFAF 60
Db 61 VAFHPASTWKNLDFKFTGY- LKDYSEVSESKDYKRLVFEFSKMGLYDKKH- IMPATIC 118
QY 61 IAYHPGTAWRHLDPLFTGYIYLKDFEVSISKDYRRLNEMSRSGIFERKKGHHIMW-TFV 119
Db 119 FIAMLFAMSVYGVLCFEGVHLFSGCLMGFLWIOSGWIHGDAGHYVMVYSDRLNKFMI 178
QY 120 GVAVMAAIYGVGLASESVGVHMLCGLLGLLWIAQAAVYGHDSGHYQVMPTRGYNRI 179
Db 179 FAANCLSGISIGWKNWNAHIAACNSLEYDPDLOYPVAVSTRFNSITSVFYGRVLK 238
QY 180 IAGNLTGISIAWKTWNAHHLACNSLDYDPDLOHPIVFAVSTRFNSITSVFYGRVLK 239
Db 239 FDSLRFVSYQHTWTFYPTMCAARLNMYVQSLMLLTNRNYSYRAQELLCGLVFSIWPL 298
QY 240 FDEVARFLVSYQHTWTFYPMIFGRVNLFIQTFTLLLTREDVDRALNLMGIAVFWT 299
Db 299 LVSCLPNNGERIMFVIALSVTGMQOVQVSLNHFSSSVYVGPKNNGWPEKOTDGTGLD 358
QY 300 FVSCLPNWPREFGVLIISFAVTAIQHVQFTLNHFSGDTYVGPKNNGWPEKOTKGTID 359
Db 359 CPPWMDWFFGGQFOIEHLLFPKMPRCNLRKISPYVIELCKHNLPI-NYASFSKANEM 417
QY 360 CPPWMDWFFGGQFOLEHLLFPRLPGOLRKIAPLARDLCKKHGMPYRSGFGWDDANVT 419
Db 418 LRTLNTALQARDITK-PLPKNLVW-EALHTHG 448
QY 420 IRTLDAVAQVARDLNSAPCPKKLGYGEAYNTHG 452

RESULT 2
ID US-08-789-936-5 STANDARD; PRT; 448 AA.
XX
AC xxxxxx
DE

Sequence 5, Application US/08789936
CC      SEQUENCE 5, Application US/08789936
CC      Patent No. 5789220
CC      GENERAL INFORMATION:
CC      APPLICANT: Thomas, Terry L.
CC      APPLICANT: Reddy, Avutu S.
CC      APPLICANT: Nuccio, Michael
CC      APPLICANT: Freyinet, Georges L.
CC      APPLICANT: Nunberg, Andrew N.
CC      TITLE OF INVENTION: PRODUCTION OF GAMMA LINOLENIC ACID BY A
CC      NUMBER OF SEQUENCES: 25
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE: Scully, Scott, Murphy & Presser
CC      STREET: 400 Garden City Plaza
CC      CITY: Garden City
CC      STATE: New York
CC      COUNTRY: United States
CC      ZIP: 11530
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

CC      SOFTWARE: PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/08/789,936
CC      FILING DATE: 28-JAN-1997
CC      CLASSIFICATION: 435
CC      PRIOR APPLICATION DATA:
CC      PRIOR APPLICATION NUMBER: 08/366,779
CC      FILING DATE: 30-DEC-1994
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: Presser, Leopold
CC      REGISTRATION NUMBER: 19,827
CC      REFERENCE/DOCKET NUMBER: 83832YXW
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: (516) 742-4343
CC      TELEFAX: (516) 742-4366
CC      TELEX: 230 901 SANS UR
CC      INFORMATION FOR SEQ ID NO: 5:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 448 amino acids
CC      TYPE: amino acid
CC      TOPOLOGY: linear
CC      MOLECULE TYPE:
SQ      SEQUENCE 448 AA: 51599 MW: 1110962 CN;

Query Match
Best Local Similarity 60.3%; Score 2121; DB 2; Length 448;
Matches 261; Conservative 105; Mismatches 81; Indels 6; Gaps 6;

Db 1 MAAQIKKTYITDELKNDKPGDLWISIOGKAYDVSDWKDHPGGSPFLKSLAGQEVTDFAF 60
QY 1 MEGEAKKITYITADLRHNKSGDLWISIOGKAYDVSDWKDHPGGSPFLKSLAGQEVTDFAF 60
Db 61 VAFHPASTWKNLDFKFTGY- LKDYSEVSESKDYKRLVFEFSKMGLYDKKH- IMPATIC 118
QY 61 IAYHPGTAWRHLDPLFTGYIYLKDFEVSISKDYRRLNEMSRSGIFERKKGHHIMW-TFV 119
Db 119 FIAMLFAMSVYGVLCFEGVHLFSGCLMGFLWIOSGWIHGDAGHYVMVYSDRLNKFMI 178
QY 120 GVAVMAAIYGVGLASESVGVHMLCGLLGLLWIAQAAVYGHDSGHYQVMPTRGYNRI 179
Db 179 FAANCLSGISIGWKNWNAHIAACNSLEYDPDLOYPVAVSTRFNSITSVFYGRVLK 238
QY 180 IAGNLTGISIAWKTWNAHHLACNSLDYDPDLOHPIVFAVSTRFNSITSVFYGRVLK 239
Db 239 FDSLRFVSYQHTWTFYPTMCAARLNMYVQSLMLLTNRNYSYRAQELLCGLVFSIWPL 298
QY 240 FDEVARFLVSYQHTWTFYPMIFGRVNLFIQTFTLLLTREDVDRALNLMGIAVFWT 299
Db 299 LVSCLPNNGERIMFVIALSVTGMQOVQVSLNHFSSSVYVGPKNNGWPEKOTDGTGLD 358
QY 300 FVSCLPNWPREFGVLIISFAVTAIQHVQFTLNHFSGDTYVGPKNNGWPEKOTKGTID 359
Db 359 CPPWMDWFFGGQFOIEHLLFPKMPRCNLRKISPYVIELCKHNLPI-NYASFSKANEM 417
QY 360 CPPWMDWFFGGQFOLEHLLFPRLPGOLRKIAPLARDLCKKHGMPYRSGFGWDDANVT 419
Db 418 LRTLNTALQARDITK-PLPKNLVW-EALHTHG 448
QY 420 IRTLDAVAQVARDLNSAPCPKKLGYGEAYNTHG 452

RESULT 3
ID US-08-801-972-1 STANDARD; PRT; 155 AA.
XX
AC xxxxxx
DE

Sequence 1, Application US/08801972
CC      SEQUENCE 1, Application US/08801972
CC      Patent No. 5831018
CC      GENERAL INFORMATION:

```

CC	STATE:	CA
CC	COUNTRY:	USA
CC	ZIP:	94304
CC	COMPUTER READABLE FORM:	
CC	MEDIUM TYPE:	Diskette
CC	COMPUTER:	IBM Compatible
CC	OPERATING SYSTEM:	DOS
CC	SOFTWARE:	FastSeq for Windows Version 2.0
CC	CURRENT APPLICATION DATA:	
CC	APPLICATION NUMBER:	US/08/801,972
CC	FILING DATE:	Filed Herewith
CC	PRIOR APPLICATION DATA:	
CC	APPLICATION NUMBER:	
CC	FILING DATE:	
CC	ATTORNEY/AGENT INFORMATION:	
CC	NAME:	Billings, Lucy J.
CC	REGISTRATION NUMBER:	36,749
CC	REFERENCE/DOCKET NUMBER:	PF-0216 US
CC	TELECOMMUNICATION INFORMATION:	
CC	TELEPHONE:	415-855-0555
CC	TELEFAX:	415-845-4166
CC	INFORMATION FOR SEQ ID NO:	4:
CC	SEQUENCE CHARACTERISTICS:	
CC	LENGTH:	134 amino acids
CC	TYPE:	amino acid
CC	STRANDEDNESS:	single
CC	TOPOLOGY:	linear
CC	IMMEDIATE SOURCE:	
CC	LIBRARY:	Genbank
CC	CLONE:	476415
SQ	SEQUENCE	134 AA; 15349 MW; 97278 CN;
	Query Match	4.9%; Score 171; DB 2: Length 134;
	Best Local Similarity	40.0%; Pred.No. 1.03e-04;
	Matches	28; Conservative 11; Mismatches 31; Indels 0; Gaps
Db	10 KYTLEETIKHNHKSWSLIIHHKYDITKLEEHPGGEVIREQGADATENFEDVGHS	69
QY	7 KYIATDLRRHKSGDLWISIQGVDCSRWAAEHPGVEPLLSLAGQDVTDFAFYHPG	66
Db	70 TDARELSKTF	79
QY	67 TAWRHLDPLF	76
RESULT	5	
ID	US-08-801-972-3	STANDARD; PRT; 142 AA.
XX	xxxxxx	
DE	Sequence 3, Application US/08801972	
CC	Sequence 3, Application US/08801972	
CC	Patent No. 5831018	
CC	GENERAL INFORMATION:	
CC	APPLICANT:	Goli, Surya K.
CC	TITLE OF INVENTION:	A NOVEL HUMAN CYTOCHROME
CC	TITLE OF INVENTION:	B5
CC	NUMBER OF SEQUENCES:	5
CC	CORRESPONDENCE ADDRESS:	
CC	ADDRESSEE:	Incyte Pharmaceuticals, Inc.
CC	STREET:	3174 Porter Drive
CC	CITY:	Palo Alto
CC	STATE:	CA
CC	COUNTRY:	USA
CC	ZIP:	94304
CC	COMPUTER READABLE FORM:	
CC	MEDIUM TYPE:	Diskette
CC	COMPUTER:	IBM Compatible
CC	OPERATING SYSTEM:	DOS
CC	SOFTWARE:	FastSeq for Windows Version 2.0

```
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Presser, Leopold  
CC REGISTRATION NUMBER: 19,827  
CC REFERENCE/DOCKET NUMBER: 838ZYYXW  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (516) 742-4343  
CC TELEXFAX: (516) 742-4366  
CC TELETYPE: 230 901 SANS UR  
CC INFORMATION FOR SEQ ID NO: 2:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 359 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
SQ SEQUENCE 359 AA; 41425 MW; 698417 CN;  
  
Query Match      4.4%; Score 155; DB 1; Length 359;  
Best Local Similarity 32.8%; Pred. No. 1.67e-03;  
Matches    22; Conservative   16; Mismatches 26; Indels 3; Gaps 3;  
  
Db 273 DEWAICQIRITANPNPFW-WNFCGGLNHQVTHHLFPNICHIHYPLENIKDVCOEF 331  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
QY 345 DNFWEKTGTGTIDI-TCPF-WMDWFEGLGQLFEHLLFLPRLPQGOLRKIAPLARDICKKH 402  
  
Db 332 GVYEKYV 338  
| : | : |  
QY 403 GMPYRSF 409  
  
RESULT       7  
ID US-08-789-936-2 STANDARD; PRT; 359 AA.  
XX xxxxxx  
AC  
XC  
AD  
DT  
XX  
DE  
CE  
CC Sequence 2, Application US/08789936  
CC Patent No. 5789220  
CC GENERAL INFORMATION:  
CC APPLICANT: Thomas, Terry L.  
CC APPLICANT: Reddy, Avutu S.  
CC APPLICANT: Nuccio, Michael  
CC APPLICANT: Freysinet, Georges L.  
CC APPLICANT: Nunberg, Andrew N.  
CC TITLE OF INVENTION: PRODUCTION OF GAMMA LINOLENIC ACID BY A  
CC TITLE OF INVENTION: DELTA 6-DESATURASE  
CC NUMBER OF SEQUENCES: 25  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Scully, Scott, Murphy & Presser  
CC STREET: 400 Garden City Plaza  
CC CITY: Garden City  
CC STATE: New York  
CC COUNTRY: United States  
CC ZIP: 11530  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent In Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/789,936  
CC FILING DATE: 28-JAN-1997  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: 08/366,779  
CC FILING DATE: 30-DEC-1994  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Presser, Leopold  
CC REGISTRATION NUMBER: 19,827  
CC REFERENCE/DOCKET NUMBER: 838ZYYXW  
CC TELECOMMUNICATION INFORMATION:
```

CC TELEPHONE: (516) 742-4343
CC TELEFAX: (516) 742-4366
CC TELEX: 230 901 SANS UR
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 359 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 359 AA; 41425 MW; 698417 CN;

Query Match 4.4%; Score 155; DB 2; Length 359;
Best Local Similarity 32.8%; Pred. No. 1.67e-03;
Matches 22; Conservative 16; Mismatches 26; Indels 3; Gaps 3;

Db 273 DEWAICQIRTTANFATNPFV-NWFCGGLNHQVTHLFPNICHYHYPOLNIIKDVCOEF 331
QY 345 DNWFEKQTKGTIDI-TCPP-WMDWFFGGLQFOLEHHLFRLPRGQLRKIAPLARDLCKKH 402
Db 332 GVEYKVV 338
QY 403 GMPYRSF 409

RESULT 8
ID US-08-307-382-2 STANDARD; PRT; 359 AA.
XX
AC xxxxxx
XX
DT
XX
DE Sequence 2, Application US/08307382
XX
CC Sequence 2, Application US/08307382
CC Patent No. 5552306
CC GENERAL INFORMATION:
CC APPLICANT: Thomas, Terry L.
CC APPLICANT: Reddy, Avutu S.
CC APPLICANT: Nuccio, Michael
CC APPLICANT: Freyssinet, Georges L.
CC TITLE OF INVENTION: PRODUCTION OF GAMMA LINOLENIC
CC TITLE OF INVENTION: ACID BY A DELTA 6-DESATURASE
CC NUMBER OF SEQUENCES: 3
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Scully, Scott, Murphy & Presser
CC STREET: 400 Garden City Plaza
CC CITY: Garden City
CC STATE: New York
CC COUNTRY: United States
CC ZIP: 11530
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/307,382
CC FILING DATE: 14-SEP-1994
CC CLASSIFICATION: 800
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/817,919
CC FILING DATE: 08-JAN-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: McNulty, William E.
CC REGISTRATION NUMBER: 22,606
CC REFERENCE/DOCKET NUMBER: 83832
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (516) 742-4343
CC TELEFAX: (516) 742-4366
CC TELEX: 230 901 SANS UR
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 359 amino acids

CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 359 AA; 41425 MW; 698417 CN;

Query Match 4.4%; Score 155; DB 1; Length 359;
Best Local Similarity 32.8%; Pred. No. 1.67e-03;
Matches 22; Conservative 16; Mismatches 26; Indels 3; Gaps 3;

Db 273 DEWAICQIRTTANFATNPFV-NWFCGGLNHQVTHLFPNICHYHYPOLNIIKDVCOEF 331
QY 345 DNWFEKQTKGTIDI-TCPP-WMDWFFGGLQFOLEHHLFRLPRGQLRKIAPLARDLCKKH 402
Db 332 GVEYKVV 338
QY 403 GMPYRSF 409

RESULT 9
ID US-08-478-727-2 STANDARD; PRT; 359 AA.
XX
AC xxxxxx
XX
DT
XX
DE Sequence 2, Application US/08478727
XX
CC Sequence 2, Application US/08478727
CC Patent No. 5653068
CC GENERAL INFORMATION:
CC APPLICANT: Thomas, Terry L.
CC APPLICANT: Reddy, Avutu S.
CC APPLICANT: Nuccio, Michael
CC APPLICANT: Freyssinet, Georges L.
CC TITLE OF INVENTION: PRODUCTION OF GAMMA LINOLENIC
CC TITLE OF INVENTION: ACID BY A DELTA 6-DESATURASE
CC NUMBER OF SEQUENCES: 3
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Scully, Scott, Murphy & Presser
CC STREET: 400 Garden City Plaza
CC CITY: Garden City
CC STATE: New York
CC COUNTRY: United States
CC ZIP: 11530
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/478,727
CC FILING DATE: 07-JUN-1995
CC CLASSIFICATION: 800
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/307,382
CC FILING DATE: 14-SEP-1994
CC APPLICATION NUMBER: US 07/817,919
CC FILING DATE: 08-JAN-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: McNulty, William E.
CC REGISTRATION NUMBER: 22,606
CC REFERENCE/DOCKET NUMBER: 83832
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (516) 742-4343
CC TELEFAX: (516) 742-4366
CC TELEX: 230 901 SANS UR
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 359 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 359 AA; 41425 MW; 698417 CN;

Query Match 4.4%; Score 155; DB 1; Length 359;
Best Local Similarity 32.8%; Pred. No. 1.67e-03;
Matches 22; Conservative 16; Mismatches 26; Indels 3; Gaps 3;

Db 273 DEWAICQIRTTANFNPFW-NWFCGGLNHQVTHLFPNICHHPQLENIKDYCQEF 331
QY 345 DNWFEKQTGTIDI-TCPP-WMDWFGGLQFQLEHLLFRLPRGQLKTIAPLARDLCKKH 402

Db 332 GVEYKVY 338
QY 403 GMPYRSF 409

RESULT 10
ID US-08-473-508-2 STANDARD; PRT: 359 AA.
XX
AC xxxxxx
XX

Sequence 2, Application US/08473508

Sequence 2, Application US/08473508
Patent No. 5689050
GENERAL INFORMATION:
APPLICANT: Thomas, Terry L.
APPLICANT: Reddy, Avutu S.
APPLICANT: Nuccio, Michael
APPLICANT: Freyssinet, Georges L.
TITLE OF INVENTION: PRODUCTION OF GAMMA LINOLENIC
TITLE OF INVENTION: ACID BY A DELTA 6-DESATURASE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/473,508
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/307,382
FILING DATE: 14-SEP-1994
APPLICATION NUMBER: US 07/817,919
FILING DATE: 08-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: McNulty, William E.
REGISTRATION NUMBER: 22,606
REFERENCE/DOCKET NUMBER: 8383Z
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 359 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE 359 AA; 41425 MW; 698417 CN;

Query Match 4.4%; Score 155; DB 1; Length 359;
Best Local Similarity 32.8%; Pred. No. 1.67e-03;
Matches 22; Conservative 16; Mismatches 26; Indels 3; Gaps 3;

Db 273 DEWAICQIRTTANFNPFW-NWFCGGLNHQVTHLFPNICHHPQLENIKDYCQEF 331
QY 345 DNWFEKQTGTIDI-TCPP-WMDWFGGLQFQLEHLLFRLPRGQLKTIAPLARDLCKKH 402

Db 332 GVEYKVY 338
QY 403 GMPYRSF 409

RESULT 11
ID US-08-801-972-5 STANDARD; PRT: 104 AA.
XX
AC xxxxxx
XX
DT XX

Sequence 5, Application US/08801972

Sequence 5, Application US/08801972
Patent No. 5831018
GENERAL INFORMATION:
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: A NOVEL HUMAN CYTOCHROME
TITLE OF INVENTION: B5
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,972
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0216 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 104 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1217655
SEQUENCE 104 AA; 11672 MW; 58506 CN;

Query Match 3.9%; Score 137; DB 2; Length 104;
Best Local Similarity 37.9%; Pred. No. 3.57e-02;
Matches 25; Conservative 11; Mismatches 29; Indels 1; Gaps 1;

Db 1 GRVYDITRFLSEHPGEEVLLQAGADATESFDYGHSPDAREMLKQYYIGDVHPNDLKP 60
QY 29 GRVYDCSRWAAEHGPGVEPLLSLAGQDVDTAFIAY-HPGTANRHLDPLFTGYVYLKDFEV 87

Db 61 KGDGXD 66
QY 88 SEISKD 93

```

XX Sequence 2, Application US/08276099A
XX
XX Sequence 2, Application US/08276099A
XX Patent No. 5591825
XX GENERAL INFORMATION:
XX APPLICANT: McNight, Steven L
XX APPLICANT: Hou, Jinhao
XX TITLE OF INVENTION: INTERLEUKIN-4 SIGNAL TRANSDUCERS AND
XX TITLE OF INVENTION: BINDING ASSAYS
XX NUMBER OF SEQUENCES: 17
XX CORRESPONDENCE ADDRESS:
XX ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
XX STREET: 4 Embarcadero Center, Suite 3400
XX CITY: San Francisco
XX STATE: California
XX COUNTRY: USA
XX ZIP: 94111-4187
XX COMPUTER READABLE FORM:
XX MEDIUM TYPE: Floppy disk
XX COMPUTER: IBM PC compatible
XX OPERATING SYSTEM: PC-DOS/MS-DOS
XX SOFTWARE: PatentIn Release #1.0, Version #1.25
XX CURRENT APPLICATION DATA:
XX APPLICATION NUMBER: US/08/276,099A
XX FILING DATE: 15-JUL-1994
XX CLASSIFICATION: 435
XX ATTORNEY/AGENT INFORMATION:
XX NAME: Osman, Richard Aron
XX REGISTRATION NUMBER: 36,627
XX REFERENCE/DOCKET NUMBER: A-59451-1/RAO
XX TELECOMMUNICATION INFORMATION:
XX TELEPHONE: (415) 781-1989
XX TELEFAX: (415) 398-3249
XX TELEX: 910 277299
XX INFORMATION FOR SEQ ID NO: 2:
XX SEQUENCE CHARACTERISTICS:
XX LENGTH: 847 amino acids
XX TYPE: amino acid
XX TOPOLOGY: linear
XX MOLECULE TYPE: protein
XX SEQUENCE 847 AA; 94134 MW; 3802900 CN;
XX
XX Query Match 2.9%; Score 102; DB 1; Length 847;
XX Best Local Similarity 21.1%; Pred.No. 1.01e+01;
XX Matches 24; Conservative 36; Mismatches 44; Indels 10; Gaps 9;
XX
Db 433 SEMDRV-FVVAERVPPWEKMCETLNLFKMAEVTGNRGLLP-EHFLFLAQKIFNDNSLSME 490
QY 211 PDLGHIPVFVAVSTR-L-FNSITSVFGRVYKLFDEVARFLVSYQHWTYYPPWIFGRVNLFIQ.269
: : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 491 AFQURSVSWSGFNKEI-LLGRGFTFWQFDGVLDTTKRCLRSYWSDRLLIGFIS 543
: | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
QY 270 TFEILLTRDPDRALNLMGIA-VFWTWFP--L-FVS-CLPN-WPERFGFVLIS 317
: : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
RESULT 14
ID US-08-624-125-6 STANDARD; PRT; 175 AA.
XX
XX xxxxxx
XX
XX
XX
XX
XX Sequence 6, Application US/08624125
XX
XX Sequence 6, Application US/08624125
XX Patent No. 5744341
XX GENERAL INFORMATION:
XX APPLICANT: CUNNINGHAM JR., FRANCIS X.
XX APPLICANT: SUN, ZAIREN
XX TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND
XX TITLE OF INVENTION: METABOLISM AND A SYSTEM FOR SCREENING SUCH GENES
XX NUMBER OF SEQUENCES: 21

```


CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202

REF: 22002
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/624,125
 FILING DATE: 29-MAR-1996

FILING DATE: 29-MAR-1996
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: KELBER, STEVEN B.
 REGISTRATION NUMBER: 30,073
 REFERENCE/DOCKET NUMBER: 2747-063-27
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-413-3000

TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 175 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE 175 AA: 19867 MW: 149

Query Match	2.8%;	Score 97;	DB 1;	Length 175;
Best Local Similarity	28.8%;	Pred. No.	2.17e+01;	
Matches	23;	Conservative	22;	Mismatches 30;
				Indels 5;
				Gaps 5;

37 HHTRPRKGVFELNDLFAV-VFAGVAITALIAGTACGWPLOQWICGMYVGLLYFL-VHDGL 94
: : | : | : : | : | : : | : | : : | : | : : :
99 NEMSRSGIFELKGGHHIMWTFVGAVNMAAI-VYGLVASESGVHMLCGALIGLLWIONAY 157

```

95 V-HQRWPFHWIPRRGYLKRL 113
    | | : : : | | | : | :
158 VGHDSCGHYQVMPTRGY-NRI 176

```

RESULT 15
US-08-096-623A-18 STANDARD; PRT: 176 AA.

XXXXXX

XXXXXX

APPLICANT: Ausich, Rodney L.
 APPLICANT: Brikhaus, Friedhelm L.
 APPLICANT: Mukharji, Indrani
 APPLICANT: Proffitt, John H.
 APPLICANT: Yarger, James G.
 APPLICANT: Yen, Hwei-Che B.
 TITLE OF INVENTION: Biosynthesis of
 TITLE OF INVENTION: Glycosylated Z
 NUMBER OF SEQUENCES: 104

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Welsh & Katz, Ltd.
 STREET: 120 S. Riverside Plaza, 22nd Floor
 CITY: Chicago
 STATE: IL
 COUNTRY: USA

```

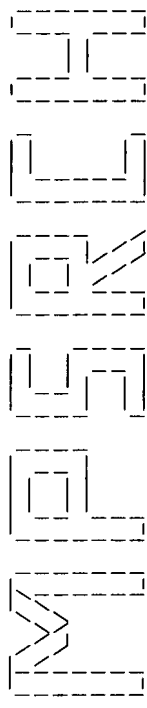
CC ZIP: 60606
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/096,623A
CC FILING DATE: 22-JUL-1993
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/805,061
CC FILING DATE: 09-DEC-1991
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/662,921
CC FILING DATE: 28-FEB-1991
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/562,674
CC FILING DATE: 03-AUG-1990
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/525,551
CC FILING DATE: 18-MAY-1990
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/487,613
CC FILING DATE: 02-MAR-1990
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Gamson, Edward P.
CC REGISTRATION NUMBER: 29,381
CC REFERENCE/DOCKET NUMBER: AMO-006.1
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (312) 655-1500
CC TELEFAX: (312) 655-1501
CC INFORMATION FOR SEQ ID NO: 18:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 176 amino acids
CC TYPE: amino acid
CC STRANDEDNESS:
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 176 AA: 19990 MW: 151305 CN:

```

Query Match 2.8%; Score 97; DB 1; Length 176;
Best Local Similarity 28.8%; pred. No. 2.17e+01;
Matches 23; Conservative 22; Mismatches 30; Indels

Db	38	HHTPRKGVFELNDFAV-VFAGVAIALIAVGTGAVNPLQWIGCMVYGLLYL-VHDGL	95
QY	99	NEMSRGIFEKKHHMTVEGVAVMMAAI-VYGVLAASESVGHMLCGALLGLWTAAY	157
Db	96	V-HQRPFPHWIPRRGYIKRL	114
QY	158	VGHDSGHYQVMPTRGY-NRI	176

Search completed: Thu Feb 18 11:38:43 1999
Job time : 27 secs.



Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Thu Feb 18 11:39:01 1999; MacPar time 23.20 Seconds
733.938 Million cell updates/sec
Linear output not generated.

Title: >US-08-934-254-27
Description: (1-452) from US08934254.pap
Perfect Score: 3515
Sequence: 1 MEGAKKYITAEDLRHNSK.....LNSAPCPKRLGYGEYNTFHG 452

Scoring table: PAM 150
Gap 11

Searched: 321224 seqs, 37675139 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-pending
1:P9 2:U60 3:U7 4:U80 5:U81 6:U82 7:U83 8:U84A 9:U84B
10:U85 11:U86 12:U87 13:U88 14:U89 15:U90 16:U91 17:U92
18:NEWP 19:NEWU8 20:NEWU9

Statistics: Mean 36.598; Variance 169.128; scale 0.216
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES					Pred. No.	
Rank	Score	Query Match	Length	ID	Description	
1	3515	100.0	452	14	US-08-934-	Sequence 27, Applicati
2	2433	89.2	458	16	US-08-116-	Sequence 5, Applicatio
3	2133	60.7	448	2	US-60-110-	Sequence 5, Applicatio
4	2121	60.3	448	14	US-08-934-	Sequence 5, Applicatio
5	2108	60.0	446	13	US-08-834-	Sequence 15, Applicati
6	2108	60.0	445	14	US-08-956-	Sequence 16, Applicati
7	2108	60.0	446	13	US-08-833-	Sequence 5, Applicatio
8	2108	60.0	446	14	US-08-956-	Sequence 16, Applicati
9	1561	44.4	393	2	US-60-110-	Sequence 4, Applicatio
10	1557	44.3	252	14	US-08-956-	Sequence 9, Applicatio
11	1557	44.3	252	14	US-08-956-	Sequence 9, Applicatio
12	1557	44.3	252	13	US-08-834-	Sequence 8, Applicatio
13	1557	44.3	252	13	US-08-834-	Sequence 8, Applicatio
14	1457	41.5	286	2	US-60-110-	Sequence 9, Applicatio
15	1025	29.2	253	2	US-60-110-	Sequence 11, Applicati
16	897	25.5	162	2	US-60-110-	Sequence 13, Applicati
17	795	22.6	202	2	US-60-110-	Sequence 2, Applicatio
18	713	20.3	125	13	US-08-834-	Sequence 8, Applicatio
19	713	20.3	125	13	US-08-834-	Sequence 9, Applicatio
20	713	20.3	125	14	US-08-956-	Sequence 10, Applicati
21	713	20.3	125	14	US-08-956-	Sequence 10, Applicati

SQ SEQUENCE 452 AA; 51763 MW; 1074264 CN;

Query Match 100.0%; Score 3515; DB 14; Length 452;
 Best Local Similarity 100.0%; Pred. No. 0.00e+00;
 Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MEGAKEYITAEEDLRHNSGDLWISIOGKYVDCSRWAAEHPGGEVPLLSLAGQDVTDAF 60
 Qy 1 MEGAKEYITAEEDLRHNSGDLWISIOGKYVDCSRWAAEHPGGEVPLLSLAGQDVTDAF 60
 Db 61 IAYHPGTAWRHLDPLFTGYIYKDFEVESEISKDYRRLNEMSRGIFFEKKGHHINWTFVG 120
 Qy 61 IAYHPGTAWRHLDPLFTGYIYKDFEVESEISKDYRRLNEMSRGIFFEKKGHHINWTFVG 120
 Db 121 VAVMAAIVYGLASESGVHMLCAGLLGWLWIAQAAVYGHDSGHYQVMPTRGYNRITOLI 180
 Qy 121 VAVMAAIVYGLASESGVHMLCAGLLGWLWIAQAAVYGHDSGHYQVMPTRGYNRITOLI 180
 Db 181 AGNLTGSIAMWKTNAHHLACNSLDYDPDLOHIPPFAVSTRLFNSTITSFYGRVLKF 240
 Qy 181 AGNLTGSIAMWKTNAHHLACNSLDYDPDLOHIPPFAVSTRLFNSTITSFYGRVLKF 240
 Db 241 DEVARFLSYQHWIYYPVMIIFGRVNLFTOTFLLLLTRDVPDRALNMGIAVFWTFWPLF 300
 Qy 241 DEVARFLSYQHWIYYPVMIIFGRVNLFTOTFLLLLTRDVPDRALNMGIAVFWTFWPLF 300
 Db 301 VSCLPNPERFGVLIISFAVTAIOHVQFTLNHFSGDTYVGPCKGDNWFEKOTKTIDITC 360
 Qy 301 VSCLPNPERFGVLIISFAVTAIOHVQFTLNHFSGDTYVGPCKGDNWFEKOTKTIDITC 360
 Db 361 PPMDWFFGGLOFQLEHLLFRLPRGQLRKIAPLARDLCKKHGMPYRSGFWDANVTI 420
 Qy 361 PPMDWFFGGLOFQLEHLLFRLPRGQLRKIAPLARDLCKKHGMPYRSGFWDANVTI 420
 Db 421 RTLDDAAVOARDLNSAPCKPKLGYEAYNTHG 452
 Qy 421 RTLDDAAVOARDLNSAPCKPKLGYEAYNTHG 452

RESULT 2
 ID US-09-116-639-5 STANDARD; PRT; 458 AA.
 XX
 AC xxxxxx
 XX
 DT
 XX
 XX
 Sequence 5, Application US/09116639
 Sequence 5, Application US/09116639
 GENERAL INFORMATION:
 APPLICANT: Bandman, Olga
 APPLICANT: Hillman, Jennifer L.
 APPLICANT: Gorgone, Gina
 APPLICANT: Guebler, Karl J.
 APPLICANT: Corley, Neil C.
 APPLICANT: Baughn, Mariah R.
 TITLE OF INVENTION: HUMAN CYTOCHROMES
 CURRENT APPLICATION NUMBER: US/09/116,639
 CURRENT FILING DATE: 1998-07-16
 NUMBER OF SEQ ID NOS: 6
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 5
 LENGTH: 458
 TYPE: PRT
 ORGANISM: HELIANTHUS ANNUUS
 FEATURE:
 OTHER INFORMATION: 1040729, GenBank
 SQ SEQUENCE 458 AA; 52231 MW; 1118207 CN;

Query Match 69.2%; Score 2433; DB 16; Length 458;
 Best Local Similarity 66.7%; Pred. No. 1.16e-208;
 Matches 300; Conservative 92; Mismatches 54; Indels 4; Gaps 4;

Db 13 ADCKYITSKELKNNPNDLWISILGKYVNTWAKHEHPGGDAPLINLAGQDVTDAFIA 72
 Qy 3 GEAKYITAEEDLRHNSGDLWISIOGKYVDCSRWAAEHPGGEVPLLSLAGQDVTDAFIA 62
 Db 73 PHPGTAWRHLDPLFTGYH-LKDYQVSDISRDYRKLADEFKAGMEKKGHHINWTFVGA 131
 Qy 63 YHPGTAWRHLDPLFTGYIYKDFEVESEISKDYRRLNEMSRGIFFEKKGHHINWTFVGA 122
 Db 132 LLLSACYGVLYSGFWIHLMSGAILGLAWMOIAYLGHDAHYQMMATRGWKNKFAFIFIG 191
 Qy 123 VMMAAIVYGLASESGVHMLCAGLLGWLWIAQAAVYGHDSGHYQVMPTRGYNRITOLI 182
 Db 192 NCITGISTANWKTNAHHLACNSLDYDPDLOHLPMLAVSSKLFNSITSVFYGRQITFDP 251
 Qy 183 NILTGISTANWKTNAHHLACNSLDYDPDLOHIPPFAVSTRLFNSTITSFYGRVLKFE 242
 Db 252 LARFVSVQHYLYYIMCVARVNLVQLTLLLSKRIKIPDRGINILGLIFWTFWPLFVS 311
 Qy 243 VARFLVSYQHWIYYPVMIIFGRVNLFTOTFLLLLTRDVPDRALNMGIAVFWTFWPLFVS 302
 Db 312 RLPNMPERVAFLVFCVTGHIQFTLNHFSGDVVYVGPCKGDNWFEKOTKTIDITACSS 371
 Qy 303 CLPNPREFGVLIISFAVTAIOHVQFTLNHFSGDTYVGPCKGDNWFEKOTKTIDITC 362
 Db 372 WMDWFFGGLOFQLEHLLFRLPRCHLRSIPICRELCKKYNLPYVLSLSPYD-ANVTTLKT 430
 Qy 363 WMDWFFGGLOFQLEHLLFRLPRGQLRKIAPLARDLCKKHGMPYRSGFWDANVTI 422
 Db 431 LETAALQARDL-TNPAPONLAW-EAFNTHG 458
 Qy 423 LRDAAVOARDLNSAPCKPKLGYEAYNTHG 452

RESULT 3
 ID US-60-110-784-5 STANDARD; PRT; 448 AA.
 XX
 AC xxxxxx
 XX
 DT
 XX
 XX
 Sequence 5, Application US/60110784
 Sequence 5, Application US/60110784
 GENERAL INFORMATION:
 APPLICANT: Kinney, Anthony J.
 APPLICANT: Cahoon, Edgar B.
 APPLICANT: Hitz, William D.
 APPLICANT: Cahoon, Rebecca E.
 TITLE OF INVENTION: MEMBRANE-BOUND DESATURASES
 FILE REFERENCE: BB-1264-P1
 CURRENT APPLICATION NUMBER: US/60/110,784
 CURRENT FILING DATE: 1998-12-03
 NUMBER OF SEQ ID NOS: 17
 SOFTWARE: Microsoft Word Version 7.0A
 SEQ ID NO 5
 LENGTH: 448
 TYPE: PRT
 ORGANISM: Borago officinalis
 SQ SEQUENCE 448 AA; 51634 MW; 1109964 CN;

Query Match 60.7%; Score 2133; DB 2; Length 448;
 Best Local Similarity 57.8%; Pred. No. 7.28e-181;
 Matches 262; Conservative 105; Mismatches 80; Indels 6; Gaps 6;

Db 1 MAAQIKYITSDKLNHDKPGDLWISIOGKAYDVSDVWVDKHPGSGFPLKSLAGQEVTD 60
 Qy 1 MEGAKEYITAEEDLRHNSGDLWISIOGKYVDCSRWAAEHPGGEVPLLSLAGQDVTDAF 60
 Db 61 VAFHPASTWKNLDKFTGYI-LKDYSEVSEVSKDYRKLVFEFSKMGLYDKKHG-INPATIC 118
 Qy 61 IAYHPGTAWRHLDPLFTGYIYKDFEVESEISKDYRRLNEMSRGIFFEKKGHHINWTFV 119

Query Match
60.3%; Score 2121; DB 14; Length 448;

```

CC ZIP: 94306
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible.
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/956,985
CC FILING DATE: 24-OCT-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/834,033
CC FILING DATE: 11-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/833,610
CC FILING DATE: 11-APR-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: RAE-VENTER, BARBARA
CC REGISTRATION NUMBER: 32,750
CC REFERENCE/DOCKET NUMBER: CGNE.128.01US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (650) 328-4400
CC TELEFAX: (650) 328-4477
CC TELEX: N/A
CC INFORMATION FOR SEQ ID NO: 16:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 446 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: not relevant
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC SEQUENCE 446 AA; 51415 MW; 1102977 CN;

Query Match 60.0%; Score 2108; DB 14; Length 446;
Best Local Similarity 57.8%; Pred. No. 1.50e-178;
Matches 236; Conservative 104; Mismatches 78; Indels 5; Gaps

Db 1 MAAQIKYITSDLELNKDKPGDLWISIQGAYDVSDWKDPGSGPPLKSLAQEVTDAF 60
QY 1 MEGAKEYITAEIDLRRHNKSGDLWISIQGKVDYCSRWAAEHGPGVPLLSLAQGDVDAF 60

Db 61 VAFHPASTWNLKDFYGY- LKDYSEVSEKVKRLKLVFEESKMGLYDKKGH- IMFATLC 118
QY 61 IAYHPGTAWRHLDPLFTGYTYLKDFFEYSEISKDYRLLNEMWSRGIFEKKGHIMW-TFV 119

Db 119 FIAMLFAMSYGVGLFCGVLVHLHFGCLMGFLMTQSGWIGHDAGHYMVVSDSLNKFMGI 178
QY 120 GVAYMMAAIVGVVLASDSVGHMLCGALLGLLQIAAYVGHDSGHVQVMPTRYNRTQL 179

Db 179 FAANCLSGISIGWKNHNAHACNLESDYDPLQIYIPFLVSKSGFGSLTSHFYEKRLT 238
QY 180 IAGNIIIGISIAWKKWTHNAHLACNSLDYDPLQIHPIYFAVSTRLENSITSVFYGEVLK 239

Db 239 FDSLRSFVSYQHWTFYPIMCAARLNMYVOSLIMLTTRKNVSYRAQELGCLVFSIYPL 298
QY 240 FDEVARELVSYQHWTYIPYPMVIFGRVNLFIQTLELLLTRDPDPRALNLMGIAVFTWFL 358

Db 299 LVSCLPNWRERIMFVIASLVTGMQVQFSLNHFSSSVYVGKPKGNWFEKQTDGLDIS 358
QY 300 FVSCLPNWRERFGLVLSFAVTAIQHYQFTLNHFSGDYYVGPKPKGNWFEKQTKGIDIT 359

Db 359 CPPWMDWFHGGFOIEHHLLFPKMPRCNLRKISPYVELCKKNLNP- NYAFSEKANEMT 417
QY 360 CPPWMDWFHGGFOIEHHLLFPKMPRCNLRKISPYVELCKKNLNP- NYAFSEKANEMT 419

Db 418 IRLTNTALQARDITK-PLPKNL 439
QY 420 IRLTDAVAQARDLNSAPCPKKL 442

RESULT 7
ID US-08-833-610-5 STANDARD: PRT; 446 AA.
XX

```

AC	XXXXXX
DB	Sequence 5, Application US/08833610
DE	Sequence 5, Application US/08833610
DT	GENERAL INFORMATION:
DX	APPLICANT: KUTZON, DEBORAH
XX	APPLICANT: MURKJJI, PRADIP
XX	APPLICANT: HUANG, YUNG-SHENG
XX	APPLICANT: THURMOND, JENNIFER
XX	APPLICANT: CHAUDHARY, SUNITA
CC	TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS
CC	TITLE OF INVENTION: OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS
CC	NUMBER OF SEQUENCES: 12
CC	CORRESPONDENCE ADDRESS:
CC	ADDRESSEE: RAE-VENTER LAW GROUP, P.C.
CC	STREET: 260 SHERIDAN AVE, P.O. BOX 60039
CC	CITY: PALO ALTO
CC	STATE: CALIFORNIA
CC	COUNTRY: USA
CC	ZIP: 94306
CC	COMPUTER READABLE FORM:
CC	MEDIUM TYPE: Floppy disk
CC	COMPUTER: IBM PC compatible
CC	OPERATING SYSTEM: PC-DOS/MS-DOS
CC	SOFTWARE: PatentIn Release #1.0, Version #1.30
CC	CURRENT APPLICATION DATA:
CC	APPLICATION NUMBER: US/08/833,610
CC	FILING DATE: 11-APR-1997
CC	CLASSIFICATION: 435
CC	ATTORNEY/AGENT INFORMATION:
CC	NAME: RAE-VENTER, BARBARA
CC	REGISTRATION NUMBER: 32,750
CC	REFERENCE/DOCKET NUMBER: CGNE.123.00US
CC	TELEPHONE: (650)328-4400
CC	TELEFAX: (650)328-4477
CC	TELEX: N/A
CC	INFORMATION FOR SEQ ID NO: 5:
CC	SEQUENCE CHARACTERISTICS:
CC	LENGTH: 446 amino acids
CC	TYPE: amino acid
CC	STRANDEDNESS: not relevant
CC	TOPOLOGY: linear
CC	MOLECULE TYPE: peptide
CC	SEQUENCE 446 AA; 51415 MW; 1102977 CN;
CC	Query Match 60.0%; Score 2108; DB 13; Length 446;
CC	Best Local Similarity 57.8%; Pred. No. 1.50e-178;
CC	Matches 256; Conservative 104; Mismatches 78; Indels 5; Gaps 5;
DB	1 MAQAKIYITSDLKNHDKPGDLWISIOGKAYDVSDWKDHGGSGFPPLKSLAGQEVTDAF 60
QY	1 MEGAKKIYITAEARRHNKSGDLWISIOGVYVSRWAAEHGPGEVPLLSLAGQDVTDFA 60
DB	61 VAFHPASTWNKLDKFTGY-LKDYVSSEVSKYVRKLVEFSKMGLYDKKGH-IMEATIC 118
QY	61 IAYHPGTAWRLDPFTGYIKDFEVEISKDYRRLLNEMSGSIFEKKGHIMW-TFV 119
DB	119 FIAMLFAMSVYGVLFCGVLVHLFGSCGLMGFLWGSIQGWIGHDAGHYMVVSDSLRNKFMI 178
QY	120 GVAVMMAIVYGLASESVGVHMLCGALLGLLIQAAYVGHDGSHYQVMPTRGYNRIQL 179
DB	179 FAANCLSGISGWKNNNAHHIACNSLEYDDPDIQYIPLVYSKFFGSLSHFYEKRUT 238
QY	180 IAGILITGISIAWKWTNAHHLACNSLDYDQLQHVPVFAVSTRLENITSVFYGRVLK 239
DB	239 FDSLRSRFVSYQHTFYDIMCAARLNMYVOSLIMLLTKRNVSYRAQELGCLVFSIWYPL 298
QY	240 FDEVARFLVSYQHTFYPMVIGFRVNLFIQTLELLLTTRVDPRALNLMGIADVFWTFWPL 299

QY 1 MEGEAKKYITAEDLRHNSGDLWISIOGKVYDCSRWAAEHGGEVPLLSLAGQDVTDAF 60
 Db 61 VAFHPASTWKNLDKFTGYG-LKDYVSSESVKVRKLVFEFSKMGYDKKGGH-IMEATLC 118
 QY 61 IAYHPGTAWRHLDPLFTGYG-LKDFEVSSEISDYRLLNEMSRSGIFERKKGHHIMW-TFV 119
 Db 119 FIAMLFAMSVYGVLCFEGVLHFLFSGCLMGFLWISQSGWIGHDAGHYMYVSDSLNFKFMGI 178
 QY 120 GVAVMAAIVYGVLAASESVGVHMLCGALLGLLWIOAAYVGHDSGHVQVMPTRGYNRITQL 179
 Db 179 FAANLCSGISGKWNHNAHNAHNAHNAHNAHNAHNAHNAHNAHNAHNAHNAHNAHNAHNAH 238
 QY 180 IAGNLTGSIAMWKTWTHNAHNAHNAHNAHNAHNAHNAHNAHNAHNAHNAHNAHNAHNAHNA 239
 Db 239 FDSLSRFFSVYOHWTFTYPTMCAARLNMVYQSLMLITLTKRNVSYRAQELLCGLVFSIWP 298
 QY 240 FDEVARFLSVYOHWTFTYPTMCAARLNMVYQSLMLITLTKRNVSYRAQELLCGLVFSIWP 299
 Db 299 LVSCLPNWGERIMFVTSLSVGMQOVQSLNHFSSSVYVGVKPNWFEKOTDGLDLS 358
 QY 300 FVSCLPNWPERGFLVLSFAVTAIHQVFTLNHFSGDTVGPVKGNWFEKOTDGLDLS 359
 Db 359 CPPWMDWFGGLOFQIEHLEFLPKMPCRNKISPYVIELCKHNLFPY-NYASFSKANEMT 417
 QY 360 CPPWMDWFGGLOFQIEHLEFLPKMPCRNKISPYVIELCKHNLFPY-NYASFSKANEMT 419
 Db 418 IRLTALQARDITK-PLPKNL 439
 QY 420 IRLTALQARDITK-PLPKNL 442

RESULT 9
 ID US-60-110-784-4 STANDARD; PRT; 393 AA.

XX xxxxxx

Sequence 4, Application US/60110784

Sequence 4, Application US/60110784

GENERAL INFORMATION:
 APPLICANT: Kinney, Anthony J.
 APPLICANT: Cahoon, Edgar B.
 APPLICANT: Hitz, William D.
 APPLICANT: Cahoon, Rebecca E.
 TITLE OF INVENTION: MEMBRANE-BOUND DESATURASES
 CURRENT APPLICATION NUMBER: US/60/110,784
 CURRENT FILING DATE: 1998-12-03
 NUMBER OF SEQ ID NOS: 17
 SOFTWARE: Microsoft Word Version 7.0A
 SEQ ID NO 4
 LENGTH: 393
 TYPE: PRT
 ORGANISM: Picramnia sw
 SEQUENCE 393 AA; 45245 MW; 859480 CN;

Query Match 44.4%; Score 1561; DB 2; Length 393;
 Best Local Similarity 54.1%; Pred. No. 4,80e-128;
 Matches 204; Conservative 86; Mismatches 8; Indels 5; Gaps 5;

Db 1 ME-EPKHISQADLAHKQPGDLWISIOGKVYDCSRWAAEHGGEVPLLSLAGQDVTDAF 59
 QY 1 MEGEAKKYITAEDLRHNSGDLWISIOGKVYDCSRWAAEHGGEVPLLSLAGQDVTDAF 60
 Db 60 IAYHPGTAWRHLDPLFTGYG-LKDFEVSSEISDYRLLNEMSRSGIFERKKGHHIMW-TFV 118
 QY 61 IAYHPGTAWRHLDPLFTGYG-LKDFEVSSEISDYRLLNEMSRSGIFERKKGHHIMW-TFV 120
 Db 119 FVSCLPNWGERIMFVTSLSVGMQOVQSLNHFSSSVYVGVKPNWFEKOTDGLDLS 238
 QY 120 GVAVMAAIVYGVLAASESVGVHMLCGALLGLLWIOAAYVGHDSGHVQVMPTRGYNRITQL 239
 Db 239 FDSLSRFFSVYOHWTFTYPTMCAARLNMVYQSLMLITLTKRNVSYRAQELLCGLVFSIWP 298
 QY 240 FDEVARFLSVYOHWTFTYPTMCAARLNMVYQSLMLITLTKRNVSYRAQELLCGLVFSIWP 299

Query Match 44.3%; Score 1557; DB 14; Length 252;

Db 178 IAGNVIAGSVAAWKLDHNTTHFACNSANLDPDIOHLPITAIASPKFNSLTSYVHNCKMT 237
 QY 180 IAGNLTGSIAMWKTWTHNAHNAHNAHNAHNAHNAHNAHNAHNAHNAHNAHNAHNAHNAHNA 239
 Db 238 YDRAARFFSVYOHWTFTYPTMCAARLNMVYQSLMLITLTKRNVSYRAQELLCGLVFSIWP 297
 QY 240 FDEVARFLSVYOHWTFTYPTMCAARLNMVYQSLMLITLTKRNVSYRAQELLCGLVFSIWP 298
 Db 298 LLSRPLNPNWGERIMFVTSLSVGMQOVQSLNHFSSSVYVGVKPNWFEKOTDGLDLS 357
 QY 299 LVSCLPNWPERGFLVLSFAVTAIHQVFTLNHFSGDTVGPVKGNWFEKOTDGLDLS 358
 Db 358 TASAWDWFHGGLOFQL 374
 QY 359 TCPWMDWFGGLOFQL 375

RESULT 10
 ID US-08-956-985-9 STANDARD; PRT; 252 AA.

XX xxxxxx

Sequence 9, Application US/08956985

Sequence 9, Application US/08956985
 GENERAL INFORMATION:
 APPLICANT: KNUTZON, DEBORAH
 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS
 OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS
 NUMBER OF SEQUENCES: 28
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: RAE-VENTER LAW GROUP, P.C.
 STREET: 260 SHERIDAN AVENUE, P.O. BOX 60039
 CITY: PALO ALTO
 STATE: CA
 COUNTRY: USA
 ZIP: 94306

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/956,985
 FILING DATE: 24-OCT-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/834,033
 FILING DATE: 11-APR-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/833,610
 FILING DATE: 11-APR-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: RAE-VENTER, BARBARA
 REGISTRATION NUMBER: 32,750
 REFERENCE/DOCKET NUMBER: CGNE.128.01US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (650) 328-4400
 TELEFAX: (650) 328-4477
 TELEX: N/A
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 252 amino acids
 TYPE: amino acid
 STRANDEDNESS: not relevant
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 SEQUENCE 252 AA; 28956 MW; 353818 CN;

Query Match

[illegible]

CC INFORMATION FOR SEQ ID NO: 8:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 252 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: not relevant
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC SEQUENCE 252 AA; 28956 MW; 353818 CN;

Query Match 44.3%; Score 1557; DB 13; Length 252;
Best Local Similarity 74.6%; Pred. No. 1.12e-127;
Matches 188; Conservative 36; Mismatches 28; Indels 0; Gaps 0;

Db 1 GVLGVLAQTSVFAHQIAAALGLLWIOSAYIGHDSGHVIMSNKSYNRFQALLSGNCLT 60
Qy 127 AIVGVLAQTSVFAHQIAAALGLLWIOSAYIGHDSGHVIMSNKSYNRFQALLSGNCLT 186

Db 61 GISIAWKWTNAHHLACNSLDYDPDLQHIPVFAVSTKFFSLSLTSRFDYDRKLTGFPVARE 120
Qy 187 GISIAWKWTNAHHLACNSLDYDPDLQHIPVFAVSTKFFSLSLTSRFDYDRKLTGFPVARE 246

Db 121 LVSQHTYYPVNCGRINLFIQTLLFLSKREVDPDRALNFAGILVFTWTPFLVSCLPN 180
Qy 247 LVSQHTYYPVNCGRINLFIQTLLFLSKREVDPDRALNFAGILVFTWTPFLVSCLPN 306

Db 181 WPERFFVFTSTALQHOFTLNHFAADVVPPTGSDWFEKQAAAGTIDISCRSYMDW 240
Qy 307 WPERFFVFTSTALQHOFTLNHFAADVVPPTGSDWFEKQAAAGTIDISCRSYMDW 366

Db 241 FFGGLQFQLEHH 252
Qy 367 FFGGLQFQLEHH 378

RESULT 13
ID US-08-834-655-7 STANDARD; PRT; 252 AA.
XX AC xxxxxx
XX DT
XX DE
XX SEQUENCE 7, Application US/08834655
XX GENERAL INFORMATION:
XX APPLICANT: KUTZON, DEBORAH
XX APPLICANT: MURKERT, PRADIP
XX APPLICANT: HUANG, YUNG-SHENG
XX APPLICANT: THURMOND, JENNIFER
XX APPLICANT: CHAUDHARY, SUNITA
XX TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS
XX TITLE OF INVENTION: OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS
XX NUMBER OF SEQUENCES: 18
XX CORRESPONDENCE ADDRESS:
XX ADDRESSEE: RAE-VENTER LAW GROUP, P.C.
XX STREET: 260 SHERIDAN AVENUE, P.O. BOX 60039
XX CITY: PALO ALTO
XX STATE: CA
XX COUNTRY: USA
XX ZIP: 94306
XX COMPUTER READABLE FORM:
XX MEDIUM TYPE: Floppy disk
XX COMPUTER: IBM PC compatible
XX OPERATING SYSTEM: PC-DOS/MS-DOS
XX SOFTWARE: Patent In Release #1.0, Version #1.30
XX CURRENT APPLICATION DATA:
XX APPLICATION NUMBER: US/08/834,655
XX FILING DATE: 11-APR-1997
XX CLASSIFICATION: 435
XX ATTORNEY/AGENT INFORMATION:
XX NAME: RAE-VENTER, BARBARA
XX REGISTRATION NUMBER: 32,750
XX REFERENCE/DOCKET NUMBER: CGNE.124.00US

CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (650) 328-4400
CC TELEFAX: (650) 328-4477
CC TELEX: N/A
CC INFORMATION FOR SEQ ID NO: 7:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 252 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: not relevant
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC SEQUENCE 252 AA; 28956 MW; 353818 CN;

Query Match 44.3%; Score 1557; DB 13; Length 252;
Best Local Similarity 74.6%; Pred. No. 1.12e-127;
Matches 188; Conservative 36; Mismatches 28; Indels 0; Gaps 0;

Db 1 GVLGVLAQTSVFAHQIAAALGLLWIOSAYIGHDSGHVIMSNKSYNRFQALLSGNCLT 60
Qy 127 AIVGVLAQTSVFAHQIAAALGLLWIOSAYIGHDSGHVIMSNKSYNRFQALLSGNCLT 186

Db 61 GISIAWKWTNAHHLACNSLDYDPDLQHIPVFAVSTKFFSLSLTSRFDYDRKLTGFPVARE 120
Qy 187 GISIAWKWTNAHHLACNSLDYDPDLQHIPVFAVSTKFFSLSLTSRFDYDRKLTGFPVARE 246

Db 121 LVSQHTYYPVNCGRINLFIQTLLFLSKREVDPDRALNFAGILVFTWTPFLVSCLPN 180
Qy 247 LVSQHTYYPVNCGRINLFIQTLLFLSKREVDPDRALNFAGILVFTWTPFLVSCLPN 306

Db 181 WPERFFVFTSTALQHOFTLNHFAADVVPPTGSDWFEKQAAAGTIDISCRSYMDW 240
Qy 307 WPERFFVFTSTALQHOFTLNHFAADVVPPTGSDWFEKQAAAGTIDISCRSYMDW 366

Db 241 FFGGLQFQLEHH 252
Qy 367 FFGGLQFQLEHH 378

RESULT 14
ID US-60-110-784-9 STANDARD; PRT; 286 AA.
XX AC xxxxxx
XX DT
XX DE
XX SEQUENCE 9, Application US/60110784
XX GENERAL INFORMATION:
XX APPLICANT: Kinney, Anthony J.
XX APPLICANT: Cahoon, Edgar B.
XX APPLICANT: Hitz, William D.
XX APPLICANT: Cahoon, Rebecca E.
XX TITLE OF INVENTION: MEMBRANE-BOUND DESATURASES
XX FILE REFERENCE: BB-1264-P1
XX CURRENT APPLICATION NUMBER: US/60/110,784
XX CURRENT FILING DATE: 1998-12-03
XX NUMBER OF SEQ ID NOS: 17
XX SOFTWARE: Microsoft Word Version 7.0A
XX SEQ ID NO 9
XX LENGTH: 286
XX TYPE: PRT
XX ORGANISM: Zea mays
XX SEQUENCE 286 AA; 32723 MW; 422187 CN;

Query Match 41.5%; Score 1457; DB 2; Length 286;
Best Local Similarity 62.0%; Pred. No. 1.77e-118;
Matches 176; Conservative 58; Mismatches 49; Indels 1; Gaps 1;

Db 1 IGFWIQSGWGHDSGHHRITGHPLLDVVQVLSNCLTGLSIWKNHNTHTHACNSL 60
Qy 148 LGLLWIOAAYVGHDSGHVIMSNKSYNRFQALLSGNCLTGLSIWKNHNTHTHACNSL 207



(TW)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Feb 18 11:35:18 1999; MasPar time 18.91 Seconds
895.315 Million cell updates/sec
Similar output not generated.

Title: >US-08-934-254-27
Description: (1-452) from US08934254.pep
Perfect Score: 3515
Sequence: 1 MEGEAKYITAEDLRHNKS.....LNSAPCPKILGYGEAYNTHG 452

Scoring table: PAM 150
Gap 11

Searched: 116738 seqs, 37463448 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir58
1:pir1 2:pir2 3:pir3 4:pir4

Statistic: Mean 49.791; Variance 105.274; scale 0.473

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	2433	69.2	458	2	S68358		hypothetical protein	0.00e+00
2	220	6.3	884	2	S66308		substrate-inducible n	4.72e-17
3	219	6.2	881	2	S25445		nitrate reductase (NA	6.84e-17
4	217	6.2	917	2	S35228		nitrate reductase (NA	1.44e-16
5	209	5.9	139	2	S4306		cytochrome b5 - commo	2.74e-15
6	203	5.8	120	2	S63052		cytochrome b5 - yeast	2.44e-14
7	203	5.8	137	2	S46307		cytochrome b5 - rice	2.44e-14
8	203	5.8	900	2	S47029		nitrate reductase (NA	2.44e-14
9	205	5.8	917	1	RDW0NH		nitrate reductase (NA	1.18e-14
10	200	5.7	911	1	RDTONH		nitrate reductase (NA	7.25e-14
11	202	5.7	918	2	A41667		nitrate reductase (NA	3.51e-14
12	199	5.7	926	1	RDSPNH		nitrate reductase (NA	1.04e-13
13	196	5.6	898	1	RDJNH		nitrate reductase (NA	3.07e-13
14	195	5.5	135	2	S49200		cytochrome b5 - commo	4.39e-13
15	193	5.5	135	1	CBRT5M		cytochrome b5, outer	9.00e-13
16	193	5.5	427	2	G70590		probable desA3 protei	9.00e-13
17	195	5.5	880	2	S72503		substrate-inducible n	4.39e-13
18	195	5.5	880	2	S66307		substrate-inducible n	4.39e-13
19	194	5.5	904	1	RDNTNT		nitrate reductase (NA	6.29e-13
20	193	5.5	904	1	RDNTNS		nitrate reductase (NA	9.00e-13
21	193	5.5	909	2	JDN065		nitrate reductase (NA	9.00e-13
22	192	5.5	982	2	S16292		nitrate reductase (NA	1.29e-12
23	189	5.4	368	2	S54809		linoleoyl-CoA desatur	3.75e-12

24	189	5.4	891	1	RDBHNP	nitrate reductase (NA	3.75e-12
25	188	5.3	74	2	A41085	nitrate reductase (NA	5.34e-12
26	187	5.3	501	2	S51160	nitrate reductase (NA	7.61e-12
27	186	5.3	912	1	RDBHNS	nitrate reductase (NA	1.08e-11
28	186	5.3	915	1	RDBHNS	nitrate reductase (NA	1.08e-11
29	183	5.2	617	2	S19254	nitrate reductase (NA	4.44e-11
30	183	5.2	916	2	S07554	nitrate reductase (NA	3.12e-11
31	182	5.2	920	2	S52301	nitrate reductase (NA	4.44e-11
32	174	5.0	87	2	S07959	cytochrome b5, hepati	7.20e-10
33	176	5.0	905	2	JN0803	nitrate reductase (NA	3.60e-10
34	171	4.9	98	2	JN0316	cytochrome b5, solubl	2.02e-09
35	172	4.9	133	2	S07964	cytochrome b5, hepati	1.43e-09
36	171	4.9	134	1	CBR85	cytochrome b5, hepati	2.02e-09
37	171	4.9	134	1	CBR05	cytochrome b5 - horse	2.02e-09
38	170	4.8	97	2	D24211	cytochrome b5, erythr	2.85e-09
39	170	4.8	98	1	CBH05E	cytochrome b5, erythr	2.85e-09
40	170	4.8	98	2	JC5783	soluble cytochrome b5	2.85e-09
41	170	4.8	133	2	C24211	cytochrome b5, hepati	2.85e-09
42	170	4.8	134	1	CBH05	cytochrome b5, micros	2.85e-09
43	170	4.8	134	2	JC5782	membrane-bound cytoch	2.85e-09
44	167	4.8	134	1	CBR05	cytochrome b5, micros	7.93e-09
45	168	4.8	147	2	S67453	hypothetical protein	5.64e-09

ALIGNMENTS

RESULT 1
ENTRY #type complete
TITLE hypothetical protein - common sunflower
ORGANISM #formal_name Helianthus annuus #common_name common sunflower
DATE 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 10-Sep-1997
ACCESSIONS S68358
REFERENCE S68358
#authors Sperling, P.; Schmidt, H.; Heinz, E.
#journal Eur. J. Biochem. (1995) 232:798-805
#title A cytochrome-b(5)-containing fusion protein similar to plant acyl lipid desaturases.
#accession S68358
#status preliminary
#molecule_type mRNA
#residues 1-458 #label SPE
#cross-references EMBL:X87143; NID:g1040728; PID:g1040729
CLASSIFICATION #superfamily cytochrome b5 core homology
FEATURE
16-90 #domain cytochrome b5 core homology #label CB5
SUMMARY #length 458 #molecular-weight 52231 #checksum 8719

Query Match	69.2%	Score	2433	DB 2	Length	458	
Best Local Similarity	66.7%	Pred. No.	0.00e+00				
Matches	300	Conservative	92	Mismatches	4	Gaps	4
Db	13	ADGKKYITSKELAKHNPN	DLWISILGKYVNTWAKEHPGGDAPLINLAGQDVTDAFTA	72			
QY	3	GEAKKYITAEEDLRHNK	SGDLWISIOGKYDCSRWAAEHPPGGEVPLSLAGQDVTDAFTA	62			
Db	73	FHPGTAKWLHDKLFTGYH-L	KDYQVSDISRDYRKLA	SEFAKAGMEPKKGHGYISLCFYS	131		
QY	63	YHPGTARHLDPLFTGY	YLLKDFEISEIKDYRRLNEMSRSGIFEKKGHHIHWTFVGVA	122			
Db	132	LLHSACVYGYLVSGSWI	HMLSGAIIGLAWMQIAYLGHDAGHYQMMATRGWKNKFAFIFIG	191			
QY	123	VMAAAIVYGVLA	SESVGVHMLCGALLGLLWIAQAAVYGHDSGHVQVMPTRGYNRITQIAG	182			
Db	192	NCITGISIAWKWTHNAH	IAACNSLDYDPDLQHLPLMVA	SSKLFNSITSVFYGRQLTFDP	251		
QY	183	NILTGISIAWKWTHNAH	LAACNSLDYDPDLQHIPVFAVSTRLFNSITSVFYGRVLKFE	242			
Db	252	LARFFVSYQHYLYP	IMCVARVNLQITILLISKRKIPDRGLNIGLITFWTWPLLVYS	311			
QY	243	VARFLVSYQHYLYP	IMCVARVNLFTOTELLTLTRDDVPDRALNMGIAVFWTFWPLFVS	302			
Db	312	RLPNPVARVFLV	SCVTGIQIOTFLNHFSGDVYVGGPKGNWFEKQTRGTIDTACSS	371			

[illegible]

```
KEYWORDS      molybdopterin-binding domain homology
               chromoprotein; electron transfer; FAD; flavoprotein; heme;
               homodimer; iron; metalloprotein; molybdenum; molybdopterin;
               NAD; nitrate assimilation; oxidoreductase

FEATURE
  93-485      #domain molybdopterin-binding domain homology #label
               PMO\
  545-619      #domain cytochrome b5 core homology #label CB5\
  667-917      #domain cytochrome-b5 reductase homology #label CBR\
  197          #binding_site molybdopterin (Cys) (covalent) #status
               predicted\
  436          #disulfide_bonds interchain #status predicted\
  580,603      #binding_site heme iron (His) (axial ligands) #status
               predicted\
  731,889      #binding_site NAD (Lys, Cys) #status predicted\
  771          #binding_site FAD (Tyr) #status predicted
SUMMARY      #length 917 #molecular-weight 103125 #checksum 2177

Query Match      6.2%  Score 217; DB 2; Length 917;
Best Local Similarity 38.4%; Pred. No. 1.44e-16;
Matches 28; Conservative 18; Mismatches 26; Indels 1; Gaps 1;

541 MNTASKMYTISE-VRKHTVADSAAIIHGHVYDCTFLKDPHGGTDSILINACTDCTEEF 599
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1 MEGEAKKIYTAEDLRHNRKSGDLWISIQGVYDCSRWAAEHGPGVEPILLSLAGQVDYDAF 60

Db 600 EATHSDKAKKLE 612
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 IAYHPGTAWRHLD 73

RESULT 5
ENTRY      S46306      #type complete
TITLE      cytochrome b5 - common tobacco
ORGANISM   #formal_name Nicotiana tabacum #common_name common tobacco
DATE       27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change
               08-Sep-1997

ACCESSIONS S46306; S33157
REFERENCE   S46306
#authors   Smith, M.A.; Stobart, A.K.; Shewry, P.R.; Napier, J.A.
#journal   Plant Mol. Biol. (1994) 25:527-537
#title     Tobacco cytochrome b(5): cDNA isolation, expression analysis
               and in vitro protein targeting.
#accession S46306
##status   preliminary
##molecule_type mRNA
##residues 1-139 #label SMI
##cross-references EMBL:X71441; NID:g296385; PID:g296386
CLASSIFICATION #superfamily cytochrome b5; cytochrome b5 core homology
FEATURE
  83          #domain cytochrome b5 core homology #label CB5
SUMMARY      #length 139 #molecular-weight 15336 #checksum 8533

Query Match      5.9%  Score 209; DB 2; Length 139;
Best Local Similarity 41.8%; Pred. No. 2.74e-15;
Matches 33; Conservative 17; Mismatches 27; Indels 2; Gaps 2;

Db 4 MGETKVFTELAE-VSOHNNAKCWLVTSGKVYDVTKFLDHPGSDVLLSATGKDATDDF 62
QY 1 MEGEAKKIYTAEDLRHNRKSGDLWISIQGVYDCSRWAAEHGPGVEPILLSLAGQVDYDAF 60

Db 63 EDVGHSSSARAMDLEYVVG 81
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 IAY-HPGTAWRHLDPLFTG 78

RESULT 6
ENTRY      S63052      #type complete
TITLE      cytochrome b5 - yeast (Saccharomyces cerevisiae)
ALTERNATE_NAMES protein N1949; protein YN111c
ORGANISM   #formal_name Saccharomyces cerevisiae
DATE       27-Apr-1996 #sequence_revision 03-May-1996 #text_change
               21-Aug-1998

ACCESSIONS S63052; S47919; S67347
REFERENCE   S63047
#authors   De Antoni, A.; D'Angelo, M.; Dal Pero, F.; Sartorello, F.;
               Pandolfo, D.; Lanfranchi, G.; Valle, G.
#submission submitted to the Protein Sequence Database, April 1996
#accession S63052
##molecule_type DNA
##residues 1-120 #label DEA
##cross-references EMBL:L271387; NID:gl302031; PID:e239931; PID:gl302032;
               MIPS:YN111c
##experimental_source strain S288C
REFERENCE   S47919
#authors   Truan, G.; Epinat, J.C.; Rougeulle, C.; Cullin, C.; Pompon,
               D.
#journal   Gene (1994) 149:123-127
#title     Cloning and characterization of a yeast cytochrome b
               (5)-encoding gene which suppresses ketoconazole
               hypersensitivity in a NADPH-P-450 reductase-deficient
               strain.
#accession S47919
##molecule_type DNA
##residues 1-16,'Q',18-120 #label TRU
##cross-references EMBL:L22494; NID:g431761; PID:g431762
REFERENCE   S67327
#authors   d'Antoni, A.; d'Angelo, M.; dal Pero, F.; Sartorello, F.;
               Pandolfo, D.; Lanfranchi, G.; Valle, G.
#submission submitted to the EMBL Data Library, February 1996
#description The DNA sequence of cosmid 14-13b from chromosome XIV of
               Saccharomyces cerevisiae reveals an unusually high number
               of overlapping ORFs.
#accession S67347
##molecule_type DNA
##residues 1-120 #label DAN
##cross-references EMBL:L269382; NID:gl183941; PID:e221832; PID:gl183962
GENETICS
#gene      SGD:CYB5
##cross-references SGD:S0005055; MIPS:YN111c
#map_position 14L
FUNCTION    electron transfer
CLASSIFICATION #superfamily cytochrome b5; cytochrome b5 core homology
KEYWORDS    electron transfer; heme; transmembrane protein
FEATURE
  2-77          #domain cytochrome b5 core homology #label CB5\
  102-118        #domain transmembrane #status predicted #label TMM
SUMMARY      #length 120 #molecular-weight 13297 #checksum 2403

Query Match      5.8%  Score 203; DB 2; Length 120;
Best Local Similarity 38.7%; Pred. No. 2.44e-14;
Matches 29; Conservative 19; Mismatches 25; Indels 2; Gaps 2;

Db 2 PKYV-SYQVEAHNGPENFWIIDDKYDYVSQFKDEHPGDEIIMDLGGQDATESFVDIG 60
QY 5 AKKYITAEEDLRHNRKSGDLWISIQGVYDCSRWAAEHGPGVEPILLSLAGQVDYDAFY- 63
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 HSDEALRLKGLYIG 75
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 64 HPGTAWRHLDPLFTG 78

RESULT 7
ENTRY      S46307      #type complete
TITLE      cytochrome b5 - rice
ORGANISM   #formal_name Oryza sativa #common_name rice
DATE       27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change
               08-Sep-1997
REFERENCE   S46307; S38582
#authors   Smith, M.A.; Stobart, A.K.; Shewry, P.R.; Napier, J.A.
#journal   Plant Mol. Biol. (1994) 25:527-537
#title     Tobacco cytochrome b(5): cDNA isolation, expression analysis
               and in vitro protein targeting.
#accession S46307
```

```

Qy 61 IAYHPTAWRHLDT 73

RESULT      9
ENTRY       nitrate reductase (NADH) (EC 1.6.6.1) 2 - Arabidopsis thaliana
TITLE       #formal_name Arabidopsis thaliana #common_name mouse-ear cress
ORGANISM    #sequence_revision 30-Sep-1991 #text_change 20-Mar-1998
DATE        A31821; S01641
ACCESSIONS  A31821; S01641
REFERENCE   Crawford, N.M.; Smith, M.; Bellissimo, D.; Davis, R.W. Proc. Natl. Acad. Sci. U.S.A. (1988) 85:5006-5010
#authors    Sequence and nitrate regulation of the Arabidopsis thaliana mRNA encoding nitrate reductase, a metalloflavoprotein with three functional domains.
#cross-references MUID:88276888
#accession   A31821
#molecule_type mRNA
##residues  1-917 ##label CRA
##cross-references GB:J03240; NID:g166781; PID:g166782
REFERENCE   Cheng, C.; Dewdney, J.; Nam, H.; den Boer, B.G.W.; Goodman, H.M. EMBO J. (1988) 7:3309-3314
#journal    A new locus (NIA1) in Arabidopsis thaliana encoding nitrate reductase.
#title      #cross-references MUID:89091069
#accession  S01641
#molecule_type mRNA
##residues  522-917 ##label CHE
##cross-references EMBL:X13435; NID:g16403; PID:g930002
##note      the translation of the nucleotide sequence is not complete
COMMENT     This enzyme catalyzes the reduction of nitrate to nitrite in cytoplasm; each chain contains one equivalent of FAD, heme iron, and molybdenum-pterin as prosthetic groups. This is a key enzyme involved in the first step of nitrate assimilation in plants, fungi, and bacteria.
GENETICS
#gene       NIA2
#map_position 1
COMPLEX     #superfamily nitrate reductase (NADH); cytochrome b5 core homology; cytochrome-b5 reductase homology; molybdopterin-binding domain homology; chromoprotein; electron transfer; FAD; flavoprotein; heme; homodimer; iron; metalloprotein; molybdenum; molybdopterin; NAD; nitrate assimilation; oxidoreductase
KEYWORDS
FEATURE     PCO\
            #domain molybdopterin-binding domain homology #label 87-482
            #domain cytochrome b5 core homology #label CB5\
            #domain cytochrome-b5 reductase homology #label CBR\
            #binding_site molybdopterin (Cys) (covalent) #status predicted\
            #disulfide_bonds interchain #status predicted\
            #binding_site heme iron (His) (axial ligands) #status predicted\
            #binding_site NAD (P), (Lys, Cys) #status predicted\
            #binding_site FAD (Tyr) #status predicted\
            #length 917 #molecular_weight 102844 #checksum 5640
SUMMARY
Query Match          5.8%; Score 205; DB 1; Length 917;
Best Local Similarity 38.4%; Pred. No. 1.18e-14;
Matches              28; Conservative 18; Mismatches 26; Indels 1; Gaps 1;

Db 538 MNTAKWYMSSE-VKXHNADSICWIIVGHVICTFRLMDHPGGSDILINAGTDCTEEF 596
QY 1 MEGEAKKYITAEEDLRHRNKSGDLWISIOGVKVCDSRWAAEHGPGVEPLSLAGQVDYDAF 60
DB 576 EAIHSOKAKMILE 588

#status preliminary
#molecule_type mRNA
#residues 1-137 ##label SMI
#cross-references EMBL:X75670; NID:g414704; PID:g414705
CLASSIFICATION #superfamily cytochrome b5; cytochrome b5 core homology
FEATURE 6-81 #domain cytochrome b5 core homology #label CB5
SUMMARY #length 137 #molecular_weight 15296 #checksum 3618

Query Match          5.8%; Score 203; DB 2; Length 137;
Best Local Similarity 40.5%; Pred. No. 2.44e-14;
Matches              32; Conservative 13; Mismatches 13; Indels 1; Gaps 1;

Db 1 MSNDKNKVYLEVAKHNSKDDCWLLIGKVVYNVKFLDHPGGDVLLSSTGTCDATDF 60
QY 1 MEGEAKKYITAEEDLRHRNKSGDLWISIOGVKVCDSRWAAEHGPGVEPLSLAGQVDYDAF 60
DB 61 EDVGHTTTARAMDEYYVG 79
61 IAY-HPGTAWRHLDPFTG 78

RESULT      8
ENTRY       nitrate reductase (NADH) (EC 1.6.6.1) nia - Lotus japonicus
TITLE       #formal_name Lotus japonicus
ORGANISM    #sequence_revision 19-Mar-1997 #text_change 20-Mar-1998
DATE        S47029
ACCESSIONS  S47029
REFERENCE   Waterhouse, R.N.; Smyth, A.J.; Prosser, I.M.; Forde, B.G.; Clarkson, D.T. submitted to the EMBL Data Library, July 1994
#description Cloning and characterization of the nitrate reductase gene in Lotus japonicus.
#accession  S47029
#molecule_type DNA
##residues  1-900 ##label WAT
##cross-references EMBL:X80670; NID:g517355; PID:g517356
##experimental_source strain B-129; cultivar GIFU
GENETICS
#gene       nia
#introns    327/1; 374/1; 451/3
CLASSIFICATION #superfamily nitrate reductase (NADH); cytochrome b5 core homology; cytochrome-b5 reductase homology; molybdopterin-binding domain homology; chromoprotein; electron transfer; FAD; flavoprotein; heme; iron; metalloprotein; molybdenum; molybdopterin; NAD; nitrate assimilation; oxidoreductase
KEYWORDS
FEATURE     PCO\
            #domain molybdopterin-binding domain homology #label 58-460
            #domain cytochrome b5 core homology #label CB5\
            #domain cytochrome-b5 reductase homology #label CBR\
            #binding_site molybdopterin (Cys) (covalent) #status predicted\
            #disulfide_bonds interchain #status predicted\
            #binding_site heme iron (His) (axial ligands) #status predicted\
            #binding_site NAD(P), (Lys, Cys) #status predicted\
            #binding_site FAD (Tyr) #status predicted\
            #length 900 #molecular_weight 101420 #checksum 4574
SUMMARY
Query Match          5.8%; Score 203; DB 2; Length 900;
Best Local Similarity 37.0%; Pred. No. 2.44e-14;
Matches              27; Conservative 19; Mismatches 26; Indels 1; Gaps 1;

Db 517 MNFTTKMY-SLSEVKKHNSPDGAIIIVGHVICTFRLMDHPGGADSLINAGTDCTEEF 575
QY 1 MEGEAKKYITAEEDLRHRNKSGDLWISIOGVKVCDSRWAAEHGPGVEPLSLAGQVDYDAF 60
DB 576 EAIHSOKAKMILE 588

```

Qy	61	IAYHPTAWRHLTD 73
RESULT	9	
ENTRY		RDMNH #type complete
TITLE		nitrate reductase (NADH) (EC 1.6.6.1) 2 - Arabidopsis thaliana
ORGANISM		#formal_name Arabidopsis thaliana #common_name mouse-ear cress
DATE		30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 20-Mar-1998
ACCESSIONS		A31821; S01641
REFERENCE		A31821
#authors		Crawford, N.M.; Smith, M.; Bellissimo, D.; Davis, R.W.
#journal		Proc. Natl. Acad. Sci. U.S.A. (1988) 85:5006-5010
#title		Sequence and nitrate regulation of the Arabidopsis thaliana mRNA encoding nitrate reductase, a metalloflavoprotein with three functional domains.
#cross-references		MUID:88276888
#accession		A31821
##molecule_type		mRNA
##residues		1-917 ##label CRA
##cross-references		GB:J03240; NID:g166781; PID:g166782
S01640		
#authors		Cheng, C.; Dewdney, J.; Nam, H.; den Boer, B.G.W.; Goodman, H.M.
#journal		EMBO J. (1988) 7:3309-3314
#title		A new locus (NIA1) in Arabidopsis thaliana encoding nitrate reductase.
#cross-references		MUID:89091069
#accession		S01641
##molecule_type		mRNA
##residues		522-917 ##label CHE
##cross-references		EMBL:X13435; NID:g16403; PID:g930002
##note		the translation of the nucleotide sequence is not complete
COMMENT		This enzyme catalyzes the reduction of nitrate to nitrite in cytoplasm; each chain contains one equivalent of FAD, heme iron, and molybdenum-pterin as prosthetic groups. This is a key enzyme involved in the first step of nitrate assimilation in plants, fungi, and bacteria.
GENETICS		
#gene		NIA2
#map_position		1
COMPLEX		
CLASSIFICATION		#superfamily nitrate reductase (NADH); cytochrome b5 core homology; cytochrome-b5 reductase homology; molybdopterin-binding domain homology
KEYWORDS		chromoprotein; electron transfer; FAD; flavoprotein; heme; homodimer; iron; metalloprotein; molybdenum; molybdopterin; NAD; nitrate assimilation; oxidoreductase
FEATURE		
87-482		
542-616		PON #domain cytochrome b5 core homology #label CB5\
567-917		#domain cytochrome-b5 reductase homology #label CBR\ predicted\
191		#binding_site molybdopterin (Cys) (covalent) #status predicted\
433		#disulfide_bonds interchain #status predicted\
577,600		#binding_site heme iron (His) (axial ligands) #status predicted\
731,889		#binding_site NAD (Lys, Cys) #status predicted\
771		#binding_site FAD (Tyr) #status predicted\
SUMMARY		#length 917 #molecular_weight 102844 #checksum 5640
Query Match		5.8%; Score 205; DB 1; Length 917;
Best Local Similarity		38.4%; Pred. No. 1.18e-14;
Matches		28; Conservative 18; Mismatches 26; Indels 1; Gaps 1;
Db	538	MNTAKWYMS-EVKHNASDCSWIIVGHVICTFLMDHPGGSDILINAGDTCTEEF 596
Qy	1	MEGEAKKIYTAEDLRHRNKSGDLWISIOGVKVCDSRWAAEHGPGVEPLSLAGQVDYDAF 60
Db	576	EAIHDKAKMILE 588

```

Db 597 EATHSDKAKKMLE 609
| | | | |
QY 61 IAYHPGTAWRHLD 73

RESULT 10
ENTRY RTONH #type complete
TITLE nitrate reductase (NADH) (EC 1.6.6.1) - tomato
ORGANISM #formal_name Lycopersicon esculentum #common_name tomato
DATE 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change
20-Mar-1998
JQ0373
ACCESSIONS JQ0373
REFERENCE JQ0373
#authors Daniel-Vedele, F.; Dorbe, M.F.; Caboche, M.; Rouze, P.
#journal Gene (1989) 85:371-380
#title Cloning and analysis of the tomato nitrate reductase-encoding
gene: protein domain structure and amino acid homologies in
higher plants.
#cross-references MUID:90185211
#accession JQ0373
##molecule_type DNA
##residues 1-911 ##label DAN
##cross-references GB:X14060; NID:g19282; PID:g19283
COMMENT This enzyme catalyzes the reduction of nitrate to nitrite in
cytoplasm; each chain contains one equivalent of FAD, heme iron,
and molybdenum-pterin as prosthetic groups. This is a key enzyme
involved in the first step of nitrate assimilation in plants,
fungi, and bacteria.

GENETICS nia
#gene 343/1; 390/1; 467/3
#introns
CLASSIFICATION #superfamily nitrate reductase (NADH); cytochrome b5 core
homology: cytochrome-b5 reductase homology;
molybdopterin-binding domain homology
KEYWORDS chromoprotein; electron transfer; FAD; flavoprotein; heme;
homodimer; iron; metalloprotein; molybdenum; molybdopterin;
NAD; nitrate assimilation; oxidoreductase
FEATURE 84-476
#domain molybdopterin-binding domain homology #label
PCO\
536-610 #domain cytochrome b5 core homology #label CB5\
661-911 #domain cytochrome-b5 reductase homology #label CBR\
188 #binding_site molybdopterin (Cys) (covalent) #status
predicted\
427 #disulfide_bonds interchain #status predicted\
571,594 #binding_site heme iron (His) (axial ligands) #status
predicted\
725,883 #binding_site NAD (Lys, Cys) #status predicted\
765 #binding_site FAD (Tyr) #status predicted
#length 911 #molecular-weight 102452 #checksum 3686
Query Match 5.7%; Score 200; DB 1; Length 911;
Best Local Similarity 39.7%; Pred. No. 7.25e-14;
Matches 29; Conservative 16; Mismatches 27; Indels 1; Gaps 1;

Db 532 MNTASKWMSSE-VRKHNSDSAWIIVGHYDASRFLKDPGGVDSILINAGTDCTEEF 590
| | | | |
QY 1 MEGEAKKIYTAEDLRHNSGDLWISIQGVYDCSRWAEHPGGEVPLLSLAGQDVDTAF 60
| | | | |
Db 591 DATHSDKAKKMLE 603
| | | | |
QY 61 IAYHPGTAWRHLD 73

RESULT 11
ENTRY RTONH #type complete
TITLE nitrate reductase (NADH) (EC 1.6.6.1) - winter squash
ORGANISM #formal_name Cucurbita maxima #common_name winter squash
DATE 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change
20-Mar-1998
JQ0373
ACCESSIONS JQ0373
REFERENCE JQ0373
#authors Hyde, G.E.; Crawford, N.M.; Campbell, W.H.

```

ACCESSIONS	S15959
REFERENCE	Friemann, A.; Brinkmann, K.; Hachtel, W.
#authors	Mol. Gen. Genet. (1991) 227:97-105
#journal	Sequence of a cDNA encoding the bi-specific NAD(P)H-nitrate
#title	reductase from the tree Betula pendula and identification of conserved protein regions.
#cross-references	MUID:G1260687
#accession	S15959
#molecule_type	mRNA
#residues	1-898 #label FRI
##cross-references EMBL:X54097; NID:g17924; PID:g17925	
#note	the authors translated the codon CAC for residue 31 as Arg and CGG for residue 32 as His
CLASSIFICATION	#superfamily nitrate reductase (NADR); cytochrome b5 core homology; cytochrome-b5 reductase homology;
KEYWORDS	molybdopterin-binding domain homology; chromoprotein; electron transfer; FAD; flavoprotein; heme; homodimer; iron; metalloprotein; molybdenum; molybdopterin; NAD; NADP; nitrate assimilation; oxidoreductase
FEATURE	
76-468	#domain molybdopterin-binding domain homology #label PCO\
528-602	#domain cytochrome b5 core homology #label CB5\
649-898	#domain cytochrome-b5 reductase homology #label CBR\
180	#binding_site molybdopterin (Cys) (covalent) #status predicted\
419	#disulfide_bonds interchain #status predicted\
563-586	#binding_site heme iron (His) (axial ligands) #status predicted\
713-870	#binding_site NAD(P) (Lys, Cys) #status predicted\
753	#binding_site FAD (Tyr) #status predicted
SUMMARY	#length 898 #molecular-weight 101001 #checksum 9371
Query Match	5.6%; Score 196; DB 1; Length 898;
Best Local Similarity	34.2%; Pred. No. 3,07e-13;
Matches	25; Conservative 21; Mismatches 26; Indels 1; Gaps 1;
Df	524 MNTSKFMSE-VKKHNSAESAWIIVGHYDCTHFLKDPHGADSILINAGTDCTEEF 582
QY	I : :::: I :::: I :::: I :::: I :::: I :::: I :::: I :::: I :::: I ::: 1 MEGEAKVITAEHLRRHNKSGLWISIQGVYDCSWAAEHFGGEVPLLSLAQDVTDADF 60
Df	583 DAHSDRAKMKLE 595
QY	I : I : I :
	61 IAYHPGTAWRHLD 73
RESULT	14
ENTRY	S49200 #type complete
TITLE	Cytochrome b5 - common tobacco
ORGANISM	formal_name Nicotiana tabacum #common_name common tobacco
DATE	16-Feb-1995 #sequence_revision 12-May-1995 #text_change 08-Sep-1997
ACCESSIONS	S49200; S25523
REFERENCE	Napier, J.A.; Smith, M.A.; Shewry, P.R.; Stobart, A.K.
#authors	submitted to the EMBL Data Library, July 1994
#submission	S49200
#accession	#molecule_type mRNA
#residues	##cross-references EMBL:X80008; NID:g510538; PID:g510539
TITLE	S25523
REFERENCE	Smith, M.A.; Stobart, A.K.; Shewry, P.R.; Napier, J.A.
#authors	submitted to the EMBL Data Library, August 1992
#submission	Isolation and expression pattern of a tobacco cytochrome b5
#description	cdna.
#accession	S25523
#molecule_type	mRNA
#residues	'EF' 12-14, 'O' 16-24, 'V' 26-31, 'D' 33-37, 'D' 39-43, 'D', 45-61, 'V' 63-82, 'A' 84-88, 'I' 90-94, 'N' 96-103, 105-107 ##label SMI
##cross-references EMBL:X68140; NID:g19852; PID:g19853	
CLASSIFICATION	#superfamily cytochrome b5; cytochrome b5 core homolog

QY 67 TAWRHLDPLFTGYYYLKD FVSEISKD 93

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_n n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Thu Feb 18 17:54:20 1999; MasPar time 1707.01 Seconds
1334.208 Million cell updates/sec
Near output not generated.

Title: >US-08-934-254-26
Description: (1-1702) from US08934254.seq
Perfect Score: 1702 1 CCCCCAAATTTTCATGTT.....TTTTTGTAATAAAAAAAA 1702
N.A. Sequence: Comp: GGGGTTTTTAAAGTACAA.....AAAAACCATTTTTTTTTT

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 2077918 seqs, 66906966 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: n-pending

1:P9 2:U6000 3:U6001 4:U6002A 5:U6002B 6:U6003A 7:U6003B
8:U6004A 9:U6004B 10:U6005 11:U6006 12:U6007 13:U6008A
14:U6008B 15:U6008C 16:U6009A 17:U6009B 18:U6010A
19:U6010B 20:U6011 21:U6012 22:U6013 23:U6014 24:U6015
25:U6016 26:U6017 27:U6018 28:U6019 29:U6020 30:U6021 31:U6022
32:U6023 33:U6024 34:U6025 35:U6026 36:U6027 37:U6028 38:U6029
39:U6030 40:U6031 41:U6032 42:U6033 43:U6034 44:U6035 45:U6036
46:U6037 47:U6038 48:U6039 49:U6040 50:U6041 51:U6042 52:U6043
53:U6044 54:U6045 55:U6046 56:U6047 57:U6048 58:U6049 59:U6050
60:U6051 61:U6052 62:U6053 63:U6054 64:U6055 65:U6056
66:U6057 67:U6058 68:U6059 69:U6060 70:U6061 71:U6062
72:U6063 73:U6064 74:U6065 75:U6066 76:U6067 77:U6068
78:U6069 79:U6070 80:U6071 81:U6072 82:U6073 83:U6074
84:U6075 85:U6076 86:U6077 87:U6078 88:U6079 89:U6080
90:U6081 91:U6082 92:U6083 93:U6084 94:U6085 95:U6086
96:U6087 97:U6088 98:U6089 99:U6090 100:U6091

Statistics: Mean 10.769; Variance 3.383; scale 3.183

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1702	100.0	1702	43	US-08-934- Sequence 26, Applicatio	0.00e+00
2	172	10.1	1684	39	US-08-831- Sequence 1, Applicatio	4.24e-178
3	172	10.1	1684	39	US-08-831- Sequence 1, Applicatio	4.24e-178
4	172	10.1	1685	43	US-08-934- Sequence 4, Applicatio	4.24e-178
5	116	6.8	291	14	US-60-086- Sequence 1966, Applic	2.32e-106
6	110	6.5	276	14	US-60-086- Sequence 2042, Applic	7.87e-99
7	103	6.1	267	15	US-60-089- Sequence 1341, Applic	4.23e-90
8	98	5.8	298	13	US-60-080- Sequence 1738, Applic	6.63e-84
9	98	5.8	304	15	US-60-089- Sequence 4177, Applic	6.63e-84
10	97	5.7	578	34	US-08-624- Sequence 14, Applicati	1.14e-82
11	97	5.7	578	33	US-08-569- Sequence 14, Applicati	1.14e-82
12	90	5.3	266	14	US-60-086- Sequence 3256, Applic	4.52e-74

13	91	5.3	7218	31	US-08-466- Sequence 14, Applicati	2.70e-75
14	89	5.2	578	33	US-08-569- Sequence 14, Applicati	7.53e-73
15	89	5.2	578	34	US-08-624- Sequence 14, Applicati	7.53e-73
16	86	5.1	526	15	US-60-089- Sequence 1148, Applic	3.39e-69
17	85	5.0	250	14	US-60-085- Sequence 3638, Applic	5.56e-68
18	85	5.0	250	15	US-60-089- Sequence 569, Applicat	5.56e-68
19	84	4.9	284	15	US-60-089- Sequence 503, Applicat	9.06e-67
20	84	4.9	284	13	US-60-080- Sequence 2485, Applic	9.06e-67
21	75	4.4	383	31	US-08-446- Sequence 8, Applicatio	5.87e-56
22	75	4.4	383	28	US-08-311- Sequence 8, Applicatio	5.87e-56
23	75	4.4	383	31	US-08-446- Sequence 8, Applicatio	5.87e-56
24	75	4.4	383	31	US-08-451- Sequence 8, Applicatio	5.87e-56
25	75	4.4	383	31	US-08-446- Sequence 8, Applicatio	5.87e-56
26	75	4.4	383	31	US-08-446- Sequence 8, Applicatio	5.87e-56
27	74	4.3	287	12	US-60-079- Sequence 8, Applicatio	9.08e-55
28	74	4.3	287	15	US-60-089- Sequence 574, Applic	9.08e-55
29	73	4.3	383	28	US-08-311- Sequence 8, Applicatio	1.40e-53
30	73	4.3	383	31	US-08-446- Sequence 8, Applicatio	1.40e-53
31	73	4.3	383	31	US-08-446- Sequence 8, Applicatio	1.40e-53
32	73	4.3	383	31	US-08-446- Sequence 8, Applicatio	1.40e-53
33	73	4.3	383	31	US-08-451- Sequence 8, Applicatio	1.40e-53
34	73	4.3	383	31	US-08-446- Sequence 8, Applicatio	1.40e-53
35	66	3.9	339	34	US-08-617- Sequence 8, Applicatio	2.36e-45
36	65	3.8	340	34	US-08-617- Sequence 8, Applicatio	3.44e-44
37	65	3.8	477	20	US-07-904- Sequence 10, Applicati	3.44e-44
38	65	3.8	477	20	US-07-904- Sequence 4, Applicatio	3.44e-44
39	62	3.6	340	34	US-08-617- Sequence 9, Applicatio	1.01e-40
40	62	3.6	477	20	US-07-904- Sequence 4, Applicatio	1.01e-40
41	61	3.6	477	20	US-07-904- Sequence 10, Applicati	1.42e-39
42	60	3.5	339	34	US-08-617- Sequence 8, Applicatio	1.99e-38
43	52	3.1	259	15	US-60-089- Sequence 504, Applicat	2.01e-29
44	53	3.1	392	48	US-09-045- Sequence 2, Applicatio	1.57e-30
45	53	3.1	392	48	US-09-040- Sequence 2, Applicatio	1.57e-30

ALIGNMENTS

RESULT 1

ID US-08-934-254-26 STANDARD; DNA; UNC; 1702 BP.

AC xxxxxx

DE Sequence 26, Application US/08934254

CC Sequence 26, Application US/08934254

CC GENERAL INFORMATION:

CC APPLICANT: Thomas, Terry L.

CC TITLE OF INVENTION: PRODUCTION OF GAMMA LINOLENIC ACID BY A

CC TITLE OF INVENTION: DELTA 6-DESATURASE

CC NUMBER OF SEQUENCES: 27

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Scully, Scott, Murphy & Presser

CC STREET: 400 Garden City Plaza

CC CITY: Garden City

CC STATE: New York

CC COUNTRY: United States

CC ZIP: 11530

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: PatentIn Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/934,254

CC FILING DATE:

CC CLASSIFICATION: 435

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Presser, Leopold

CC REGISTRATION NUMBER: 19,827

CC REFERENCE/DOCKET NUMBER: 83832YXWU

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: (516) 742-4343

CC TELEFAX: (516) 742-4366

CC TELEX: 230 901 SANS UR

CC INFORMATION FOR SEQ ID NO: 26:

CC REGISTRATION NUMBER: 19,827
 CC REFERENCE/DOCKET NUMBER: 8383ZYXWVU
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (516) 742-4343
 CC TELEFAX: (516) 742-4366
 CC TELEX: 230 901 SANS UR
 CC INFORMATION FOR SEQ ID NO: 4:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 1685 base pairs
 CC TYPE: nucleic acid
 CC STRANDEDNESS: both
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: DNA (genomic)
 SQ SEQUENCE 1685 BP; 431 A; 277 C; 357 G; 62

Query Match 10.1%; Score 172; DB 43; Length 1685;
Best Local Similarity 57.7%; Pred. No. 4.24e-178;
Matches 657; Conservative 0; Mismatches 507; Indels 3; Gaps 2;

Db	42	CAATGGCTGCTCAAAATCAAGAAATACATACCTACAGATGAACATCAAGAACCCACGATAAAC	101
Qy	46	CAATGGAGGCGAAGCTAAGAAGTATATACGGCGGAGGACCTCCGCGCCACAACAAGT	105
Db	102	CCGAGAGATCTATGGATCTCGATCAAGGGNAACCCATATGATGTTTCGGATTCGGGTGAAG	161
Qy	106	CCGCGCATCTCTGGATCTCATCCAGGGCAAGGTCTACGACTCTCTCGGTGGGCGGCGG	165
Db	162	ACCATCCAGGTGCGAGCTTCCCTTTGAAGAGTCTTGCTGCTCAAGAGGTAACTGATGCAT	221
Qy	166	AGCACCCGCGGCGAGGTCCCGCTCCTCAGTCTGCGCGCCAGGACGCTACCGACGCGCT	225
Db	222	TTGTTGCATTCCATCCTGCGCTACATGGAAGAAATCTTGATAAGTTATAGAAGCTGTGTTTG	281
Qy	226	TCATTGGGTACCAACCCGGGCACGGCTGGCGCATCTGGATCCGCTCTTCACCGGCTACT	285
Db	282	A-T--CTTAAAGATTACTCTGTTTCTTGAGGTTTCTTAAAGATTATAGAAGCTGTGTTTG	338
Qy	286	ACTACCTCAAGGACTTCGAAGTCTCGAGATCTCCAAGGACTACCGGAGGCTTTTGAACG	345
Db	339	AGTTTCTTAAATGGGTTTCTATGACAAAAAGGTCTATATTGTTTGAACACTTGTGCT	398
Qy	346	AGATGTCGGCTCGGGAATCTCGAAGNAAGGCCACACATCATGTGGACGCTCGTCG	405
Db	399	TTATAGCAATGCTGTTTGCTATGAGTGTTTATGGGTTTTGTTTTGTAGGGTGTTTTG	458
Qy	406	CGGTTGGGTCATGATGGCGCAATCGTCTACGCGCTGCTGGCGTCGGAGTCGTCGGAG	465
Db	459	TACATTTGTTTCTGGGTGTTTGATGGGTTTCTTTGGATTACAGATGTTGGATTGGAC	518
Qy	466	TTACATGCTCTCGCGCGCACTCTGGGCTTGCTGTGGATCCCAAGCGCGCTATGTGGGCC	525
Db	519	ATGATGCTGGGCATATATAGTAGTGTCTGATTCAAGGCTTAATAGTTTATGGGTATTT	578
Qy	526	ATGACTCGGCGCAATACCGGTGATGCCAACCGGTGGATACACAGAAATCACCCACTCA	585
Db	579	TTGCTGCAAAATGCTCTTCAGGAATAAATTTGGTTGGTGGAAATGGAACCATTAATGCAC	638
Qy	586	TAGCAGCAACATCCCTAACCGGAATCAGATCGCGTGGTGGAGTGGACCCACAACGCC	645
Db	639	ATCACATTCGCTGTAAATAGCCTTGAATATGACCGCTGATTTACAATATATACCATTCCTTG	698
Qy	646	ACCACCTTCGCTGCAACAGCCTCGACTACGACCCCGACCTCCAGCACATCCCCGTATTCG	705
Db	699	TTGTTCTTCCAAGTTTTTTGGTTCACCTACCTCTCATTTCTATGAGAAAGGTTGACTT	758
Qy	706	CCGTCTCCACCGACCTCTCAACTCCATCACCCTCGGTCTCTCTATGCCCGAGTCTGAAAT	765
Db	759	TTGACTCTTTTATCAAGATCTTTGTAAGTTATCAACATTCGGACATTTTACCCCTATTATGT	818
Qy	766	TCGACGAAGTGGCACGGTTCCTAGTCAGCTACGACGACTGGACCTACTACCCGGTCAIGA	825
Db	819	GTGCTGCTAGGCTCAATATATGTATACAATCTCTCAATATGTTGTTGACCACAGAAATG	878

Qy	826	TC	TTGGCGCGAGTCAACCTTC	TTCAATCCAGACCTTTTTTAT	TGCTCCTCACCAGGCGCGACG	885
Db	879	TG	CTCATTCAGCTCAGGAAC	CTTTGGGATGCGCTAGTGT	CTTCGATTTGGTACCGTTGC	938
Qy	886	TC	CTGACCGCGCTCAAACT	TAAATGGGTATCGCGT	TTTCTGGACGTGGTTC	945
Db	939	TT	GTTCCTGTTTGGCTAAT	ATGCGGTGAAGAAT	TATGTTTATTCGAAAGTTTATCAG	998
Qy	946	TC	GTATCTTCCTCCGAACT	GGCCTGAACGGTTCGGGT	TCGTCCTCATCAGCTTTCGGG	1005
Db	999	TG	ACTGGAATGCAACAGT	TCAGTTCTCCTTGAAACCACT	TCCTTCAAGTGTTTATGTTG	1058
Qy	1006	TC	AGCGGATCCAGCAGCT	CCAGTTCACGCTCAACCACT	TCCTCGGCGACACATACGTTG	1065
Db	1059	GA	AGCCATAAAGGGAATA	TTGGTTTGAAACAAACGAT	TGGGACACTTCGACATTTCTT	1118
Qy	1066	GC	CCCCCAGGCGCAACT	GGTTTCGAGACGACGAAAGG	GACGATCGATATCAGCT	1125
Db	1119	GT	CTCCTTGGATGGAT	TGGTTTCATGTGGATTGCAAT	TCCAAATTGAGCATCATTTGT	1178
Qy	1126	GC	CCACGCTGGATGGACT	GGTCTTTTGGTGGGCTCAGT	TCCAGTTGGAGCACCACITGT	1185
Db	1179	TT	CCAGATGCTAGATCAAC	CTTAGCAAAATCTCGCCCTAC	GTGATCGAGTTATGCA	1238
Qy	1186	TC	CTTAGCTCGCGCTGGC	AGCTTAGGAAGTTGCGCCT	TGGCTCGGACITGTGTA	1245
Db	1239	AG	AACA	1245		
Qy	1246	AG	AGCA	1252		
RESULT	5					
ID	US-60-086-722-1966	STANDARD; DNA; UNC; 291 BP.				
AC	xxxxx					
DT	Sequence 1966, Application US/60086722					
CC	Sequence 1966, Application US/60086722					
CC	GENERAL INFORMATION:					
CC	APPLICANT: Lalgudi, Raghunath					
CC	APPLICANT: Ito, Laura Y.					
CC	APPLICANT: Sherman, Bradley					
CC	TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM					
CC	TITLE OF INVENTION: CORN EARS					
CC	NUMBER OF SEQUENCES: 7600					
CC	CORRESPONDENCE ADDRESS:					
CC	ADDRESSEE: INCYTE PHARMACEUTICALS, INC.					
CC	STREET: 3174 PORTER DRIVE					
CC	CITY: PALO ALTO					
CC	STATE: CALIFORNIA					
CC	COUNTRY: USA					
CC	ZIP: 94304					
CC	COMPUTER READABLE FORM:					
CC	MEDIUM TYPE: Floppy disk					
CC	COMPUTER: IBM PC compatible					
CC	OPERATING SYSTEM: PC-DOS/MS-DOS					
CC	SOFTWARE: Word Perfect 8.0 for Windows/MS-DOS 6.2					
CC	CURRENT APPLICATION DATA:					
CC	APPLICATION NUMBER: US/60/086,722					
CC	FILING DATE: HERewith					
CC	CLASSIFICATION:					
CC	PRIOR APPLICATION DATA:					
CC	APPLICATION NUMBER:					
CC	FILING DATE:					
CC	CLASSIFICATION:					
CC	ATTORNEY/AGENT INFORMATION:					
CC	NAME: CERRONE, MICHAEL C.					
CC	REGISTRATION NUMBER: 39,132					
CC	REFERENCE/DOCKET NUMBER: PL-0017 P					
CC	TELECOMMUNICATION INFORMATION:					
CC	TELEPHONE: (650) 855-0555					
CC	TELEFAX: (650) 845-4166					
CC	INFORMATION FOR SEQ ID NO: 1966:					
CC	SEQUENCE CHARACTERISTICS:					

CC LENGTH: 291 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC IMMEDIATE SOURCE:
CC CLONE: 700551823H1
SQ SEQUENCE 291 BP; 47 A; 88 C; 80 G; 75 T; 1 OTHER.

Query Match 6.8%; Score 116; DB 14; Length 291;
Best Local Similarity 70.8%; Pred. No. 2.32e-106;
Matches 204; Conservative 0; Mismatches 83; Indels 1; Gaps 1;

Db 2 CCCGTTGCTGGGCTCCCTCCCAATTTGGTGGAGAGGTCGCGTTTGTCTTTTCAG 61
QY 938 CCGCTCTCTGATCTGTCTCCCACTGGCTGAACGGTTTCGGGTTCTGCTCATCAG 997
Db 62 CTTACCACTCTCGGGATTCAGCAGCTCCAAATCTGCCTGAACCACTTCTCTCGCGAGCT 121
QY 998 CTTTCGGTTCAGCGGATCCAGCAGCTCCAGTTCAGCTCAACCACTTCTCGGGGACAC 1057
Db 122 GTATCTCGGGCCACCCAGGGCAATGACTGTTTGAGAGCAGAGCGGACGCTCGA 181
QY 1058 ATACGTGGGCCCCCAGGGGCAACTGTTTCGAGAAGCAGAGGAGGACGATCGA 1117
Db 182 CATCTGTGCTCTCTTGGATGGATGTTCCACGTCGCTGCGCTCCAGATTGAGCA 241
QY 1118 TATCAGTGGCCACCGTGGATGGTGTCTTGTGGGCTGCACTTCCAGTTGGAGCA 1177
Db 242 CCATCTGTTCCCNCTACCTAGTGCACCTCCGTAAG-TTGGGCC 288
QY 1178 CCACCTGTTCCCTAGCTGCCCGTGGGAGCTTAGGAGATTGGCC 1225

RESULT 6

ID US-60-086-856-2042 STANDARD; DNA; UNC; 276 BP.
AC xxxxxx

DE Sequence 2042, Application US/60086856
CC Sequence 2042, Application US/60086856
CC GENERAL INFORMATION:
CC APPLICANT: Lalgudi, Raghunath
CC APPLICANT: Ito, Laura Y.
CC APPLICANT: Sherman, Bradley
CC TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
CC TITLE OF INVENTION: CORN EMBRYO
CC NUMBER OF SEQUENCES: 8690

CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
CC STREET: 3174 PORTER DRIVE
CC CITY: PALO ALTO
CC STATE: CALIFORNIA
CC COUNTRY: USA
CC ZIP: 94304

CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Word Perfect 8.0 for Windows/MS-DOS 6.2
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/60/086,856
CC FILING DATE: HEREWITH
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:

CC FILING DATE:
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: CERRONE, MICHAEL C.
CC REGISTRATION NUMBER: 39,132
CC REFERENCE/DOCKET NUMBER: PL-0016 P
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (650) 855-0555

CC TELEFAX: (650) 845-4166
CC INFORMATION FOR SEQ ID NO: 2042:
CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 276 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC IMMEDIATE SOURCE:
CC CLONE: 700261349H1
SQ SEQUENCE 276 BP; 49 A; 85 C; 74 G; 67 T; 1 OTHER.

Query Match 6.5%; Score 110; DB 14; Length 276;
Best Local Similarity 72.4%; Pred. No. 7.87e-99;
Matches 176; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

Db 3 GATTCAGCAGCTCCCAATTTGCCTGAACCACTTCTCGTCCGACGTGATGTCGGGCCACC 62
QY 1013 GATCCAGCAGCTCCAGTTCACGCTCAACCACTTCTCCGGGACACATACGTGGGCCCCC 1072
Db 63 CAAGGCGAATGACTGGTTTGAAGCAGACGCGACGCTCGACATCCTGTGCTCTCC 122
QY 1073 CAAGGCGCACTGTTTCGAGAAGCAGACGAAAGGAGCATCGATATCACGTGCCACC 1132
Db 123 TTGGATGATTTGGTTCACGCTGCGCTGCAGTTCAGATTGAGCACCATCTGTTCCCG 182
QY 1133 GTGGATGACTGGTTCCTTGGTGGCTGCAGTTCAGTTGGAGCACCATTGTTCCCTAG 1192
Db 183 CTTACTAGTTCGCCACCTCCGTAAGTTTCGCTGCGCTCCGCGAGCTTTCGAAGAAGCA 242
QY 1193 GTCGCGGTGGGAGCTTAGGAAGATTGCGCCCTGGCTCGGACTTGTGTAAAGAAGCA 1252
Db 243 TGG 245
QY 1253 CGG 1255

RESULT 7

ID US-60-089-804-1341 STANDARD; DNA; UNC; 267 BP.
AC xxxxxx

DE Sequence 1341, Application US/60089804
CC Sequence 1341, Application US/60089804
CC GENERAL INFORMATION:

CC APPLICANT: Stuve, Laura L.
CC APPLICANT: Gooding, Douglas H.
CC APPLICANT: Ito, Laura Y.
CC APPLICANT: Sherman, Bradley K.

CC TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED
CC TITLE OF INVENTION: FROM CORN EMBRYO
CC NUMBER OF SEQUENCES: 2470
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Incyte Pharmaceuticals, Inc.
CC STREET: 3174 Porter Drive
CC CITY: Palo Alto
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94304

CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette
CC COMPUTER: IBM compatible
CC OPERATING SYSTEM: DOS
CC SOFTWARE: Word Perfect 6.1/MS-DOS 6.2 software
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/60/089,804
CC FILING DATE: HEREWITH
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:

CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: CERRONE, MICHAEL M.
CC REGISTRATION NUMBER: 36,132

CC REFERENCE/DOCKET NUMBER: PL-0020 P
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 650-855-0555
CC TELEFAX: 650-845-4166
CC INFORMATION FOR SEQ ID NO: 1341:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 267 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC IMMEDIATE SOURCE:
CC CLONE: 701183619H1
CC SEQUENCE 267 BP; 48 A; 83 C; 81 G; 55 T; 0 OTHER.

Query Match 6.1%; Score 103; DB 15; Length 267;
Best Local Similarity 77.2%; Pred. No. 4.23e-90;
Matches 146; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

1 GCGAAGCACTGTTGCGAAGCAGAGCGGCGCTGAGTCCAGATTGACCATCTGTTCCCGCGCTG 60
1077 GCGCAACTGTTGCGAAGCAGAGCGGCGCTGAGTCCAGATTGACCATCTGTTCCCGCGCTG 1136

Db 61 ATGGATTGTTCCAGTGGCGCTGAGTCCAGATTGACCATCTGTTCCCGCGCTG 120
1137 ATGGATTGTTCCAGTGGCGCTGAGTCCAGATTGACCATCTGTTCCCGCGCTG 1196

Db 121 CCGCGTGGCACTTAGAAGATTGCGGCTTGGCTGGGACTTGTGTGAAGACACGGG 180
1197 CCGCGTGGCACTTAGAAGATTGCGGCTTGGCTGGGACTTGTGTGAAGACACGGG 1256

Db 181 CTGACCTAT 189
1257 ATGCGGTAT 1265

RESULT 8
ID US-60-080-506-1738 STANDARD; DNA; UNC; 298 BP.
AC xxxxxx
DT
DE Sequence 1738, Application US/60080506
CC Sequence 1738, Application US/60080506
CC GENERAL INFORMATION:
CC APPLICANT: Wu, Kunsheng
CC APPLICANT: Lalgudi, Raghunath
CC APPLICANT: Ito, Laura Y.
CC APPLICANT: Sherman, Bradley
CC TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
CC NUMBER OF SEQUENCES: 5534
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
CC STREET: 3174 PORTER DRIVE
CC CITY: PALO ALTO
CC STATE: CALIFORNIA
CC COUNTRY: USA
CC ZIP: 94304
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/60/080,506
CC FILING DATE: HEREWITH
CC CLASSIFICATION:
CC PRIOR APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: CERRONE, MICHAEL C.
CC REGISTRATION NUMBER: 39,132

CC REFERENCE/DOCKET NUMBER: PL-0007 P
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (650) 855-0555
CC TELEFAX: (650) 845-4166
CC INFORMATION FOR SEQ ID NO: 1738:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 298 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC IMMEDIATE SOURCE:
CC CLONE: 700100371H1
CC SEQUENCE 298 BP; 56 A; 90 C; 82 G; 69 T; 1 OTHER.

Query Match 5.8%; Score 98; DB 13; Length 298;
Best Local Similarity 71.7%; Pred. No. 6.63e-84;
Matches 160; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

1 CTTCTCGTCGACGTGTATGTCGGGCCACCCCAAGGCGCAATGACTGGTTTGAGAAGCAGAC 60
1043 CTTCTCGTCGACGTGTATGTCGGGCCACCCCAAGGCGCAATGACTGGTTTGAGAAGCAGAC 1102

Db 61 GCAGGCGCTGCGACATCTGCTCTCTCTGATGGATTGTTCCACGGTGGCGCTGCA 120
1103 GAAAGGCGCTGCGACATCTGCTCTCTCTGATGGATTGTTCCACGGTGGCGCTGCA 1162

Db 121 GTTCAGATTGACCATCTGTTCCCGCGCTACTCGGTGCCACTTCGCAAGCTTGC 180
1163 GTTCAGATTGACCATCTGTTCCCGCGCTACTCGGTGCCACTTCGCAAGCTTGC 1222

Db 181 ACCGCGCTCCGCGACCTTTGCAAGAAGCATGGCTCACTAT 223
1223 GCGCTTGGCTCGGACTTGTGTGAAGACACGGGATGCGGTAT 1265

RESULT 9
ID US-60-089-810-4177 STANDARD; DNA; UNC; 304 BP.
AC xxxxxx
DT
DE Sequence 4177, Application US/60089810
CC Sequence 4177, Application US/60089810
CC GENERAL INFORMATION:
CC APPLICANT: Joseph R. Byrum
CC APPLICANT: Michael D. Thompson
CC TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules
CC NUMBER OF SEQUENCES: 5409
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Lawrence M. Lavin, Jr.
CC STREET: BB4F
CC STREET: Monsanto Company
CC STREET: 700 Chesterfield Parkway North
CC CITY: St. Louis
CC STATE: MO
CC COUNTRY: United States
CC ZIP: 63198
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette (3.5in, 1.44 Mb)
CC COMPUTER: IBM PC/XT/AT, IBM PS/2 or compatibles
CC OPERATING SYSTEM: Windows 95/NT
CC SOFTWARE: Winzip 6.3
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/60/089,810
CC FILING DATE:
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Lavin Jr., Lawrence M.
CC REGISTRATION NUMBER: 30,768
CC REFERENCE/DOCKET NUMBER: 38-21(15395)A
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (314) 737-6670
CC TELEFAX: (314) 737-6047

CC CITY: EMERYVILLE
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94508
CC
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC Compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/569,578
CC FILING DATE: 08-DEC-1995
CC CLASSIFICATION: 514
CC ATTORNEY/AGENT INFORMATION:
CC NAME: GREEN, GRANT
CC REGISTRATION NUMBER: 31,259
CC REFERENCE/DOCKET NUMBER: 1182.001
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 510-601-2706
CC TELEFAX: 510-601-655-3542
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 578 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC
CC MOLECULE TYPE: protein
CC SEQUENCE 578 BP; 37 A; 11 C; 62 G; 65 T; 403 OTHER.

```
Query Match      5.7%; Score 97; DB 33; Length 578;
Best Local Similarity 12.2%; Pred. No. 1.14e-82;
Matches 70; Conservative 281; Mismatches 210; Indels 12; Gaps 11;
```

D	b	3	SARVTSRRSSRRRVGANNWWSKSTSSSDPKVA-HGSHERSVTVKASGKMVKVDVSGKKKSK	61
C	p	704	GATACGGGGAGTGCTCGAGGTCCGGGGTCGTAGTCAGCGCTGTTCCAGCGCAGGAGTGGTG	645
D	b	62	DGSDKYSHKKKNKYVTKRKRYVADSKRGRMKNHSDMTGTWNMGNAKTTSWSRGKT	121
C	p	644	GGCGTGTGGTGCCACTTCCACCACGCATGCTGATT-CGGTTAGGATGTTGCCCTGCTA	586
D	b	122	RDSADYHDYVGDTSKKKHUSTSVTKIVAHWNVRVAKHRSCHTDNVKGTANTGNKGAVG	181
C	p	585	TGAGTTGCGGTGATCTGTTGTATCCACGGGTGGCATCACTCGTAAATGCCGCGAGTCAT	526
D	b	182	VSMNGTSGVNSHTSGSKRRNNMNRAGDKSNTHRWMDNRYDVTWATKKYADSHDR	241
C	p	525	GGCCCCACATACGGCGTTGGATCCACAGCAAGCCACGACGATGCGCCGACAGCATGTGA	466
D	b	242	RKVHTTYARTRDYKAYTKATGMKNYSWCMDRVWKSYHVVCYSYGTSDMTSDHSVATAG	301
C	p	465	CTCCGA-CGGACTCCGACGCCACGACGCCGTAGACGATTCCGCCCATCATGACCGCAAG	407
D	b	302	VTSXKNG-GTVDGRCYATKTKSTKYHSSCSVK-SGNSGVVVGKFKSDY-DHKSXSDS	358
C	p	406	CGACGAAAGCTCCACATGATGTTGGTGGCCCTTCTTCGAAAGATCCCGGACCGCACAT	347
D	b	359	DSYGGCARATYTHHGTHGKTSGRTRYDYVKTRDSSGKTSTSHDMKWVTSRACSGS	418
C	p	346	TCGTTCAAAGCCTCCG-GTAGTCTTTGGAGATC-TCCGACACTTCGA-AGTCCITTAGG	290
D	b	419	TNNYMVGGMHVKTSTANSYDKDSRCRSTGSKKSTANRGTSTRSDGNKAGDTTKMY	478
C	p	289	TAGTAGTCGGGTGAAGACGGATCCATGTCGCCACGCCGTGCGCCG-GTGGTACGC	231
D	b	479	GSSSKARKDSKWRKCGSSVTKAADRGGKVARSRSTCSSAGAAGDGKSGKTVSSAVAK	538
C	p	230	AATGAAGCGTCGGTGACGCTCTGGCCGCCAGA--CTGAGGAGCGGACCTCGCCGCG	173
D	b	539	RKSRSNITTRVKSAVHHSKGRDYRNTHGKHR	571
C	p	172	GGGTGTCGCGCCGCCACCGCAGCAGTGGTAG	140

```

RESULT 12.
ID US-60-086-722-3256 STANDARD; DNA; UNC; 266 BP.
AC
xxxxx
DT
DE Sequence 3256, Application US/60086722
DE Sequence 3256, Application US/60086722
CC GENERAL INFORMATION:
CC APPLICANT: Lalgudi, Raghunath
CC APPLICANT: Ito, Laura Y.
CC APPLICANT: Sherman, Bradley
CC TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
CC TITLE OF INVENTION: CORN EARS
CC NUMBER OF SEQUENCES: 7600
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
CC STREET: 3174 PORTER DRIVE
CC CITY: PALO ALTO
CC STATE: CALIFORNIA
CC COUNTRY: USA
CC ZIP: 94304
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Word Perfect 8.0 for Windows/MS-DOS 6.2
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/60/086,722
CC FILING DATE: HERewith
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: CERRONE, MICHAEL C.
CC REGISTRATION NUMBER: 39,132
CC REFERENCE/DOCKET NUMBER: PL-0017 P
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (650) 855-0555
CC TELEFAX: (650) 845-4166
CC INFORMATION FOR SEQ ID NO: 3256:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 266 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC IMMEDIATE SOURCE:
CC CLONE: 700611205H1
CC SEQUENCE 266 BP; 47 A; 81 C; 83 G; 54 T; 1 OTHER.

Query Match 5.3%; Score 90; DB 14; Length 266;
Best Local Similarity 76.5%; Pred. No. 4,52e-74;
Matches 137; Conservative 0; Mismatches 41; Indels 1; Gaps

Db 1 GGTCGAGAAGCAGACGGCAGGACGCTCGACATCTCTGTGCCCTCTCTGGATGGATTGGT 60
QY 1087 GGTCGAGAAGCAGACGAAAGGACGATCGATACGTCGCCACCGCTGGATGGACTGGT 114
Db 61 TCCATGGCGGCTGCAGTTCACAGATTGACACCATCTGTTCGCCGCTGCTCGGTGC-C 119
QY 1147 TCATTGGTGGCTGCAGTTCACAGTTGGACACCACTGTTCCTAGGCTCCCGCGTGGCC 120
Db 120 ACCTTCGGAAGTGGCCGCGCGGTGCGTGACCTTTGCAAGAACGATGGCGCTGACCTAT 178
QY 1207 AGCTTAGGAAGATTGGCCGCTTGGCTCGGGACTTGTGTAAGAAGACACGGATCGGTAT 1565

RESULT 13
ID US-08-466-194-14 STANDARD; DNA; UNC; 7218 BP.
AC
xxxxxx
DT

```

DE Sequence 14, Application US/08466194
CC Sequence 14, Application US/08466194
CC GENERAL INFORMATION:
CC APPLICANT: DORNER, F.
CC APPLICANT: SCHEIFLINGER, F.
CC APPLICANT: FALKNER, F. G.
CC TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
CC NUMBER OF SEQUENCES: 52
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Foley & Lardner
CC STREET: 3000 K Street, NW, Suite 500
CC CITY: Washington
CC STATE: DC
CC COUNTRY: USA
CC ZIP: 20007
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/466,194
CC FILING DATE: 06-JUN-1995
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/232,463
CC FILING DATE: 22-APR-1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/935,313
CC FILING DATE: 26-AUG-1992
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: EP 91 114 300.6
CC FILING DATE: 28-AUG-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: BENT, Stephen A.
CC REGISTRATION NUMBER: 29,768
CC REFERENCE/DOCKET NUMBER: 30472/201 IMMU
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (202)672-5300
CC TELEFAX: (202)672-5399
CC TELEX: 904136
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 7218 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC CLONE: pTZ9pt-Fls
CC SEQUENCE 7218 BP; 1944 A; 1491 C; 1486 G; 1929 T; 368 OTHER.

Query Match 5.3%; Score 91; DB 31; Length 7218;
Best Local Similarity 0.3%; Pred. No. 2.70e-75;
Matches 1; Conservative 229; Mismatches 139; Indels 0; Gaps 0;

Db 1067 YY 1126
QY 634 CCACCAACGCCACCACTCGCTGCAAGAGCTGACTACGACCCGACCTCCAGCACA 693
Db 1127 YY 1186
QY 694 TCCCGGTATCGCGCTCCACCGGACTTCAACCTGATCACTGCGGTCTTATGCGC 753
Db 1187 YY 1246
QY 754 GAGTCTGAATCGAGAGTGGCAGGTCTCTAGTACGACTACGACTGACCTACT 813
Db 1247 YY 1306
QY 814 ACCCGGTATGATCTTCGCGGAGTCAACTCTCATCAGACCTTTTATGCTCTCA 873
Db 1307 YY 1366
QY 874 GZAGCGGACGCTCCCTGACCGGCTCTAAACTTAATGGGTATCGCGGTTTCTGACGT 933

Db 1367 YY 1426
QY 934 GGTTCGCGCTCTTGGTATCTGTCTCCGAAGTGGCTGACGCTGCGGTTCGCTCA 993
Db 1427 YYYYYYYG 1435
QY 994 TCAGCTTG 1002

RESULT 14
ID US-08-569-578-14 STANDARD; DNA; UNC; 578 BP.
AC xxxxx
DT

Sequence 14, Application US/08569578
Sequence 14, Application US/08569578
GENERAL INFORMATION:
CC APPLICANT: KAVANAUGH PH.D., MICHAEL
CC APPLICANT: POT PH.D., DAVID
CC APPLICANT: WILLIAMS PH.D., LEWIS T.
CC TITLE OF INVENTION: IDENTIFICATION, CLONING, AND USE OF
CC TITLE OF INVENTION: SHC-ASSOCIATED INOSITOL POLYPHOSPHATE 5-PHOSPHATASE
CC TITLE OF INVENTION: (SAPTAPE)
CC NUMBER OF SEQUENCES: 32
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: CHIRON CORPORATION
CC STREET: 4560 HORTON STREET
CC CITY: EMERYVILLE
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94608
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/569,578
CC FILING DATE: 08-DEC-1995
CC CLASSIFICATION: 514
CC ATTORNEY/AGENT INFORMATION:
CC NAME: GREEN, GRANT
CC REGISTRATION NUMBER: 31,259
CC REFERENCE/DOCKET NUMBER: 1182.001
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 510-601-2706
CC TELEFAX: 510-601-655-3542
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 578 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 578 BP; 37 A; 11 C; 62 G; 65 T; 403 OTHER.

Query Match 5.2%; Score 89; DB 33; Length 578;
Best Local Similarity 12.2%; Pred. No. 7.53e-73;
Matches 66; Conservative 259; Mismatches 207; Indels 9; Gaps 9;

Db 25 KTSDDSKVKAHSHRSVTVKASGMKVDYSGKKSKDGDGSKDKSHKKSNKVVTKRKY 84
QY 826 TCTTCGCGGAGTCAACCTCTTCATCCAGACCTTTTATGCTCCTCACAGGCGCAGC 885
Db 85 VADSKRGCMKNKHSMTGTWNMGNAKKTWSKGGKTRDSDADYH-DYVGTDSKWKHS 143
QY 886 TCCTGACCGCGCTCTAAACTTAATGGGTATCGCGGTTTCTGACGCTGTTCCCGC-TC 944
Db 144 VTKVAHTNVRVAKHNRSCTDNVKTGANTGNKGVSMNGTSGVNSHTSGSKRRNN 203
QY 945 TTGCTATCTGTCTCCGAAGTGGCTGCGGTTCGCTCCTCATCAGCTTTGCG 1004
Db 204 YNMRAGDKKSNTHRTHWGDNYRVDTWATKKYADSHDTRRKVHTATYRTRDKYATKKAT 263

Query Match 5.2%; Score 89; DB 34; Length 578;
Best Local Similarity 12.2%; Pred. No. 7.53e-73;
Matches 66; Conservative 259; Mismatches 207; Indels 9; Gaps 9;

Db 25 KSTSSDVKVKAHSGHRSYVTVKASGKMKVDVSGKSKSDGSDKYSKHKSKNKNVTKKRY 84
QY TCTTCGGCGGAGTCAACCTCTTCACGACCTTTTATGCTCTCACCAGCGCGAG 885
Db 85 VADSKKRCMKNKHSDMTGTWNMGNAKTSWKGKTRDDSDADYH-DYVGTDSKWKHS 143
QY TCCCTGACCGCGCTCTAAACTTAATGGGTATCGCGGTTTCTGGACGTGTTCCCGC-TC 944
Db 144 VTKIVAHVWNRVVAKHNRSHCTDNVKTGANTGNKGAUVSMNGTSGVNSHTSGSKRRNN 203
QY TCCCTGATCTTGTCTCCGAACTGGCTGAACGGTTTCGGGTTGCTCTCATCAGCTTTGG 1004
Db 204 YMRAGDKKSNTHRTHWGDNYRVDVWATKVDADSHDTRRKVHTATYRTRDKYATKKAT 263
QY 1005 GTCAGCGGATCCACGACGTCAGTTCAGCTCAACCACTTCTCCGGGACACATACGTG 1064
Db 264 GMKYNWCDRVWKSYYHVVCYSGTSDMTSDHSVATAGVTSVSKNGTVDSGRCYATKTK 323
QY 1065 GSCCCCGGCAAGGCGGCAACTGGTTCGAGAAGCAGAGGAGGACGATCGA-TATCAC 1123
Db 324 STKYHSSCVKSGNSGVVVKFKKSDYDHSKSSDSYGGCAR-ATTYTH-HGTGCHK 381
QY 1124 GTGCCACCGTGGATGGAGCTGTTTGGTGGGCTGAGTTCCAGTTGGAGCACCCTT 1183
Db 382 SGKTRKYDVKTRDSSG-KTKSTSHDMKWTSRACSGSSTNNYMGVGMHVKTSDTAWSYD 440
QY 1184 GTCCCTAGCTGCGGCGGACGCTTAGAAGATTGCGCCCTTGGCTCGGAGCTTTG 1243
Db 441 KDSGCRGTGSKKSTANRGRTSRSDGKNAGDTRKMYGSSSKARKDSKMRKCGSSVTKA 500
QY 1244 TAAGAAGCAGCGGATCCGCTATAGAGCTTCGGGTTTGGGACGACGCTAAT-GTCAGGA 1302
Db 501 ADRGGGKVARSTCSSAGRAAGDKSGKKTVSSAVAKRKSRSNTTTRVKSVAHVHSGRD 560
QY 1303 CAATTGGGACGCT-GAG-GGATGCGCGGTTTCAGGCGGTGACCTTAATTGCGCCCGTG 1360
Db 561 Y 561
QY 1361 C 1361

RESULT 15

ID US-08-624-190-14 STANDARD; DNA; UNC; 578 BP.

AC xxxxxx

DT

CC Sequence 14, Application US/08624190

CC Sequence 14, Application US/08624190

CC GENERAL INFORMATION:

CC APPLICANT: KAVANAUGH PH.D., MICHAEL

CC APPLICANT: POT PH.D., DAVID

CC APPLICANT: WILLIAMS PH.D., LEWIS T.

CC TITLE OF INVENTION: SIGNALLING INOSITOL POLYPHOSPHATE

CC TITLE OF INVENTION: 5-PHOSPHATASE

CC NUMBER OF SEQUENCES: 38

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: CHIRON CORPORATION

CC STREET: 4560 HORTON STREET

CC CITY: EMERYVILLE

CC STATE: CA USA

CC COUNTRY: USA

CC ZIP: 94608

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: PatentIn Release #1.0, Version #1.30

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/624,190

CC FILING DATE:

CC CLASSIFICATION: 435

CC ATTORNEY/AGENT INFORMATION:

CC NAME: GREEN, GRANT

CC REGISTRATION NUMBER: 31,259

CC REFERENCE/DOCKET NUMBER: 1182.003

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: 510-601-2706

CC TELEFAX: 510-655-3542

CC INFORMATION FOR SEQ ID NO: 14:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 578 base pairs

CC TYPE: nucleic acid

CC STRANDEDNESS: single

CC TOPOLOGY: linear

CC MOLECULE TYPE: protein

SQ SEQUENCE 578 BP; 37 A; 11 C; 62 G; 65 T; 403 OTHER.

Search completed: Thu Feb 18 18:55:18 1999
Job time : 3658 secs.

WIREH

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Thu Feb 18 16:22:44 1999; MasPar time 2188.90 Seconds
1392.438 Million cell updates/sec
Linear output not generated.

Title: >US-08-934-254-26

Description: (1-1702) from US08934254.seq

Perfect Score: 1702

N.A. Sequence: 1 CCCCAAAATTTTCATTGTT.....TTTTGGTAAAAAAAAA 1702

Comp: GGGGTTTTTAAAGTAACAA.....AAAAACCATTTTTTTTTT

Scoring table: TABLE default

Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 2275026 seqs, 895388244 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: emb1-est56

Database: genbank-est109

10:gb_est11 6:gb_est10 7:gb_est11 8:gb_est12 9:gb_est13

14:gb_est18 15:gb_est19 16:gb_est20 17:gb_est21

18:gb_est21 19:gb_est23 20:gb_est4 21:gb_est5 22:gb_est6

23:gb_est7 24:gb_est8 25:gb_est9 26:gb_gss1 27:gb_gss2

28:gb_gss3 29:gb_gss4

Statistics: Mean 11.586; Variance 2.268; scale 5.110

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Match Length DB ID Description Pred. No.

1 172 10.1 556 11 H36955 15084 Lambda-PRL2 Arab 3.96e-283

2 156 9.2 408 11 W43255 22634 Lambda-PRL2 Arab 4.38e-250

3 142 8.3 480 11 T42806 6069 Lambda-PRL2 Arabi 2.18e-221

4 127 7.5 455 11 T44542 7805 Lambda-PRL2 Arabi 6.45e-191

5 88 5.2 318 20 ATTS4723 A. thaliana transcribe 2.21e-113

6 64 3.8 562 11 N37558 18785 Lambda-PRL2 Arab 2.21e-68

7 60 3.5 252 12 AA754459 97SN1787 Rice Immature 3.94e-61

8 58 3.4 252 12 AA754459 97SN1787 Rice Immature 1.52e-57

9 55 3.2 247 12 AA754458 97SN1784 Rice Immature 3.18e-52

10 51 3.0 352 11 T42569 5832 Lambda-PRL2 Arabi 3.01e-45

11 50 2.9 597 16 T15209 crs698 Ricinus communi 1.59e-43

12 48 2.8 247 12 AA754458 97SN1784 Rice Immature 4.09e-40

13 46 2.7 2275 11 AF034173 Homo sapiens ntcon2 co 9.56e-37

14 34 2.0 448 17 A1107749 GH05586.5prime GH Dros 1.05e-17

15 34 2.0 613 17 A1107703 GH05526.5prime GH Dros 1.05e-17

16 32 1.9 470 9 AA527400 ng37b07.s1 NCI_CGAP-Co 8.75e-15

17 33 1.9 2275 11 AF034173 Homo sapiens ntcon2 co 5.12e-16

18 30 1.8 505 7 C25669 Dictyostellium discoide 5.81e-12

19 30 1.8 611 12 AA736027 GM10804.5prime GM Dros 5.81e-12

20 30 1.8 616 12 AA735895 GM10052.5prime GM Dros 5.81e-12

21 30 1.8 698 14 AA941503 LD25669.5prime LD Dros 5.81e-12

22 29 1.7 310 8 AA279374 zs84e05.s1 NCI_CGAP_GC 1.37e-10

23 29 1.7 370 12 AA767321 GH13838.5prime GH Dros 1.37e-10

24 29 1.7 564 18 A1135916 GH09329.5prime GH Dros 1.37e-10

25 29 1.7 591 17 A1109930 GH09329.5prime GH Dros 1.37e-10

26 29 1.7 833 20 W28140 44e3 Human retina cDNA 1.37e-10

27 27 1.6 100 24 AA296594 EST11122 Uterus Homo s 6.14e-08

28 28 1.6 300 9 C35646 C.elegans cDNA clone y 3.00e-09

29 30 1.6 366 28 AQ133540 HS_3051.A1.D08_MR CIT 3.00e-09

30 28 1.6 378 13 AA804782 Ob98h06.s1 NCI_CGAP_GC 3.00e-09

31 27 1.6 398 8 AA279631 zs86h09.s1 NCI_CGAP_GC 6.14e-08

32 27 1.6 420 28 AQ210456 HS_2202.B2.A08_MF CIT 6.14e-08

33 28 1.6 429 20 W53753 md13f10.r1 Soares mous 3.00e-09

34 27 1.6 452 9 AA531459 nj63c11.s1 NCI_CGAP_Pr 6.14e-08

35 27 1.6 528 10 AA642379 ns25e09.s1 NCI_CGAP_GC 6.14e-08

36 27 1.6 547 8 AA262103 zs19a05.s1 NCI_CGAP_GC 6.14e-08

37 27 1.6 565 12 C84893 Dictyostellium discoide 6.14e-08

38 28 1.6 681 22 AA148344 JM00M066.QM3 Miracidia 3.00e-09

39 28 1.6 719 26 B21675 F17P15-Sp6 IGF Arabidop 3.00e-09

40 27 1.6 752 26 B21679 F8L9-Sp6 IGF Arabidop 6.14e-08

41 27 1.6 802 18 AU035600 Mus musculus cDNA, clo 6.14e-08

42 28 1.6 809 26 B10155 F4F14-Sp6.1 IGF Arabid 3.00e-09

43 27 1.6 858 26 B18671 T29F5-Sp6 TAMU Arabidop 6.14e-08

44 27 1.6 878 26 B20048 T16G2-T7 TAMU Arabidop 6.14e-08

45 28 1.6 900 26 B08164 F11011-T7 IGF Arabidop 3.00e-09

ALIGNMENTS

1 RESULT 1 H36955 556 bp mRNA EST 30-DEC-1997

LOCUS 15084 Lambda-PRL2 Arabidopsis thaliana cDNA clone 181D127, mRNA

DEFINITION H36955 sequence.

ACCESSION H36955

NID G906454

KEYWORDS EST.

SOURCE thale cress.

ORGANISM Arabidopsis thaliana

Eukaryotae; Viridiplantae; Charophyta/Embryophyta group;

Embryophyta; Tracheophyta; seed plants; Magnoliophyta;

Magnoliopsida; Capparales; Brassicaceae; Arabidopsis.

1 (bases 1 to 556)

Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kendre,H.,

McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M.,

Rezel,E. and Somerville,C.

Genes galore: a summary of methods for accessing results from

large-scale partial sequencing of anonymous Arabidopsis cDNA clones

Plant Physiol. 106, 1241-1255 (1994)

95148729

Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University

MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing MI

Tel: 517-353-0854

Fax: 517-353-9168

Email: 22313tcn@bm.cl.msu.edu

Seq primer: T7 dye primer.

Location/Qualifiers

1..556

/organism="Arabidopsis thaliana"

/strain="var columbia"

/note="Vector: lambda Zip-Lox; Site_1: Sal; Site_2: Not;

Lambda PRL2 is a cDNA library derived from equal

quantities of 4 pools of mRNA. The mRNA sources were 1) 7

RESULT	5	ATTS4723	318 bp	RNA	EST	25-MAR-1995
LOCUS						
DEFINITION		A. thaliana transcribed sequence: clone YAY347; 5' end, mRNA				

sequence.

ACCESSION

F13728

NID

9755104

KEYWORDS

EST; expressed sequence tag; partial cDNA sequence.
thale cress.

SOURCE

ORGANISM

Arabidopsis thaliana

Eukaryotes; mitochondrial eukaryotes; Viridiplantae;

Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;

Magnoliopsida; Capparales; Brassicaceae; Arabidopsis.

1 (bases 1 to 318)

Morris, P.C., Guerrier, D., Barbet, J.C. and Giraudat, J.

Direct Submission

Submitted (24-MAR-1995) CNRS, GDR-1003 ACS, INRA, Laboratoire de

Biologie Moleculaire, BP 27, 31326 Castanet-Tolosan cedex, France.

E-mail: gdr-sv@toulouse.inra.fr. On behalf of: CNRS, ISV-CNRS, BP 1

Bat 23 91198 Gif/Yvette, France.

E-mail: giraudat@refle.isv.cnrs-gif.fr

2 (bases 1 to 318)

CNRS.

The Arabidopsis thaliana transcribed genome: the GDR cDNA program

Unpublished

Cloning vector: Lambda ZAPII non-oriented ;

Physiological conditions: greenhouse plants.

full automatic.

Location/Qualifiers

1. 318

/organism="Arabidopsis thaliana"

/db_xref="taxon:3702"

/clone="YAY347"

/tissue_type="Green siliques of A. thaliana ecotype

columbia"

/clone_lib="Gif-SeedA"

68 a 100 c 57 g 89 t 4 others

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 5.2%; Score 88; DB 20; Length 318;

Matches 189; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

Db 25 CGTCGCCGATGTCCTCGGAGTCTCTACGGTGTGTTGGTCTGCTCCCTCGCTNCCC 84

QY 407 CGTTCGGTATGATGGCGCAATCGTACGGCTGCTGGGTCGGAGTCCGTCGGAGT 466

Db 85 TCACCAATATCGCCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 144

QY 467 TCACATGCTCTGCGCGCACTGCTGGGCTGCTGGATCCCAAGCGGATGTGGGCA 526

145 CGATTCTGCTATGCTATCATGCTGCAACAAATCTAATACANATTCGCTCAGCTTCT 204

527 TGACTCCGGCATACAGGATGATGCCAACCGGTGGATACACAGATACAGCAACTCAT 586

Db 205 CTCGGTAACCTGCTACCGGAATCTAATCGCTGGTGGAAATGAGCTCACAATGCTCA 264

QY 587 AGCAGCAACATCTTACCGGAATCAGCATCGCTGGTGGAGTGGAGTGGAGTGGAGT 646

Db 265 TCATCTAGCTGTACAGCTCGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 318

QY 647 CCACCTCGCTGCAACAGCTCGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 700

RESULT 6

LOCUS N37558 562 bp mRNA EST 05-JAN-1998

DEFINITION 18785 Lambda-PRL2 Arabidopsis thaliana cDNA clone 205087, mRNA

sequence.

ACCESSION N37558

NID 91158700

KEYWORDS EST.

SOURCE thale cress.

ORGANISM Arabidopsis thaliana

Eukaryotes; Viridiplantae; Charophyta/Embryophyta group;

Embryophyta; Tracheophyta; seed plants; Magnoliophyta;

Magnoliopsida; Capparales; Brassicaceae; Arabidopsis.

1 (bases 1 to 562)

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

Newman, T., deBruijn, F.J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomas, M., Retzel, E., and Somerville, C.
Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones Plant Physiol. 106, 1241-1255 (1994)
95148729

Contact: Thomas Newman

MSU-DOE Plant Research Laboratory

Michigan State University

MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.

Lansing, MI

Tel: 517-353-0854

Fax: 517-353-9168

Email: 22313tcm@msu.edu

Seq primer: T7 dye primer.

Location/Qualifiers

1. 562

/organism="Arabidopsis thaliana"

/strain="var columbia"

/note="Vector: lambda Zip-Lox; Site_1: Sal; Site_2: Not;

Lambda PRL2 is a cDNA library derived from equal

quantities of 4 pools of mRNA. The mRNA sources were 1) 7

day germinated etiolated seedlings; 2) tissue culture

grown roots; 3) staged plants half with 24 hour light

cycle, half on 16 hr light, 8 hour dark- rosettes; 4)

same plants as 3 but aerial tissue (stems, flowers and

siliques. The vector is BRL's lambda Zip-Lox. The cDNA

inserts were directionally cloned with Sal-Not arms using

oligo dt primed cDNA."

/db_xref="taxon:3702"

/clone="205087"

/clone_lib="Lambda-PRL2"

116 a 141 c 128 g 156 t 21 others

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 3.8%; Score 64; DB 11; Length 562;

Matches 270; Conservative 0; Mismatches 197; Indels 4; Gaps 3;

Db 29 CATCCCGGAGGACACGGTATCTCAATCTGTTGGTCAAGACGTCACCGATGCTTC 88

QY 168 CACCCCGGAGGAGTCCGCTCTCAGTCTGGCGGCGGAGGACGTCACCGGCTTC 227

Db 89 ATCGCATTTTCATCCCGGAACCGCTTGGACCATCTCGACCATCTCTTCCACGGTTACAC 148

QY 228 ATTGGGTACCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 287

Db 149 -A--TCAGAGATTTCCAGTCTCCGAAGTCTCACGGGATACCGTGTATGGTCCCGAG 205

QY 288 TACCTCAAGGAGTTCGAAGTGTGGAGATCTCCAAGGACTACCGGAGGCTTTTGAACGAG 347

Db 206 TTTCGTAACCTCGGCTCTTCGAAACAAAGTCAAGTCACTCTCTACACTTAGCTTC 265

QY 348 ATGTCCGGTCCGGATCTTCGAGAAGAGGGCCACCATCATCATGTGGAGCTTCGTCGC 407

Db 266 GTCGGCGCATGTTCTCTNGGAGTCTCTACGGTGTGTTGGCTGTACCTCCGCTTCGCT 325

QY 408 GTTGGGTATGATGGCGGCAATCGTCTACGGCTGTGGCGTGTGGCGTGTGGCGAGT 467

Db 326 AACCAAAATCGCGCGGCTTCCTCGGGTCTCTCTGATGCCAGAGNGNTTACATAGGCA 385

QY 468 CACATGCTCTGGCGCACTGCTGGCT-TGCTGTGGATCCCAAGCGCGTATGTGGGCA 526

Db 386 CGTNTTGGTATACGGTATCATGTNGACAATCTNTTAAACAGTTTCGNTCAGTTT 445

QY 527 TGACTCCGCCATTACAGGTGATGCCAACCGCGGTGATACACAGAAATCACCAACTCAT 586

Db 446 TCCGGNAACTTTTACCGGATNTAATNGGGGTGGGATGCGNCA 496

QY 587 AGCAGCAACATCTTACCGGAATCAGCATCGCTGGTGGAGTGGAGTGGAGTGGAGT 637


```

RESULT 7
LOCUS AA754459 252 bp mRNA EST 20-JAN-1998
DEFINITION 97SN1787 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa
ACCESSION AA754459
NID 92801165
KEYWORDS EST.
SOURCE Rice.
ORGANISM Oryza sativa
REFERENCE 1 (bases 1 to 252)
AUTHORS Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,
Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
Lee,M.C. and Eun,M.Y.
TITLE Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
JOURNAL Unpublished (1998)
COMMENT Contact: Eun M.Y.
Department of CytoGenetics
National Inst. of Agri. Sci. and Tech, RDA
Suwon, Kyungido, Korea
Tel: 82 331 290 0301
Fax: 82 331 290 0307
Email: myeun@sun20.asti.re.kr
Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
University, Yongin, Korea. 449-728 bhna@bioserver.myongji.ac.kr
Seq primer: M13 Reverse Primer.
FEATURES
source
Location/Qualifiers
1..252
/organism="Oryza sativa"
/cultivar="Milyang23"
/Note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
XhoI; Directional cDNA library inserted into lambda ZAPII
vector at 5' end with EcoRI and 3' end with Xho I site."
/db_xref="taxon:4530"
/clone="97SN1787"
/clone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"
/tissue_type="Immature Seed"
/dev_stage="5 days after pollination"
/lab_host="E. coli SOLR"
BASE COUNT 5 a 21 c 12 g 35 t 179 others
ORIGIN
Query Match 3.5%; Score 60; DB 12; Length 252;
Best Local Similarity 11.4%; Pred. No. 3.94e-61;
Matches 25; Conservative 109; Mismatches 84; Indels 1; Gaps 1;
16 BAWMTTSYBCHGNVWVCVASHGNTMSVHNCBTBRTGTHCDCKNVNWSMTWTGTVNBNVSG 75
QY 1341 GACCTTAATCGGCCCGCTGCCTAAGAACTGGTATGGGAGCTATAACACCCAT 1400
Dy 76 DWHYWBVNTKVDGNHTRCSRWRBVTMAHYHDYTNCBYNNNDYHMBBMYBBTGC 135
QY 1401 GGTTCATTGCTGTTGTTGTTGGTGGAGGACCTCTTATTATTATGATGTCAC 1460
Dy 136 TCTMWCNHYNTKTCASGWHSTNTDVKSSNTWGTSTNWKSMHGYWCSBBVKHTKVS 195
QY 1461 AATATTGAACCTGAATACCACTGAAGGCACTACGTTACGTTAAGTTGCTTAAC 1520
Dy 196 TTRATRSYTCVRKYCVMM-TKKVVKYHVBBGCHBD 233
QY 1521 TAGCTGGTGGCTTCCTTGGTGGGGGAAAGTCAGTA 1559

RESULT 8
LOCUS AA754459 252 bp mRNA EST 20-JAN-1998
DEFINITION 97SN1787 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa
ACCESSION AA754459
NID 92801165

```

```

KEYWORDS EST.
SOURCE Rice.
ORGANISM Oryza sativa
REFERENCE 1 (bases 1 to 252)
AUTHORS Nahm,B.H., Kim,J.K., Cheong,J.J., Kim,S.I., Hahn,T.R., Moon,E.P.,
Kim,W.T., Kim,W.Y., Yang,M.S., Park,R.D., Sohn,U.I., Kang,K.Y.,
Lee,M.C. and Eun,M.Y.
TITLE Large-scale Sequencing Analysis of ESTs from Rice Immature Seed
JOURNAL Unpublished (1998)
COMMENT Contact: Eun M.Y.
Department of CytoGenetics
National Inst. of Agri. Sci. and Tech, RDA
Suwon, Kyungido, Korea
Tel: 82 331 290 0301
Fax: 82 331 290 0307
Email: myeun@sun20.asti.re.kr
Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji
University, Yongin, Korea. 449-728 bhna@bioserver.myongji.ac.kr
Seq primer: M13 Reverse Primer.
FEATURES
source
Location/Qualifiers
1..252
/organism="Oryza sativa"
/cultivar="Milyang23"
/Note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
XhoI; Directional cDNA library inserted into lambda ZAPII
vector at 5' end with EcoRI and 3' end with Xho I site."
/db_xref="taxon:4530"
/clone="97SN1787"
/clone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"
/tissue_type="Immature Seed"
/dev_stage="5 days after pollination"
/lab_host="E. coli SOLR"
BASE COUNT 5 a 21 c 12 g 35 t 179 others
ORIGIN
Query Match 3.4%; Score 58; DB 12; Length 252;
Best Local Similarity 12.7%; Pred. No. 1.52e-57;
Matches 29; Conservative 121; Mismatches 72; Indels 6; Gaps 6;
Dy 10 RGCCCBAAWMTTSYBCHGNVWVCVASHGNTMSVHNCBTBRTGTHCDCKNVNWSMTWT-WGT 67
Cp 525 GCGCCACATACGGCTGGATCCA-CAGCAAGCCAGCAGTCGCCGCGAGCATGTGA 467
Dy 68 VNBWNSGDWHYWBNTKVDGNHTRCSRWRBVTMAHYHDYTNCBYNNNDYHMBB 127
Cp 466 ACTCGGACGGACTCGGACGCGCAGCAGCGCGTAGACGATTGCGGCATCATGACCGCAAG 407
Dy 128 MYBTGCMTCMWCNHYNTKTCASGWHSTNTDVKSSNTWGTSTNWKSMHGYWCSBB 187
Cp 406 CCGAGCAAGTC-CACATGATGT-GGTGGCCCTTCTTCGAAGATCCGCGAC-CGCGAC 350
Dy 188 VKYHTKVSITRATRSYTCVRKYCVMMTKKVKYHVBBGCHBD 235
Cp 349 ATCTGTTCAAAGACCTCCGTAGTCCCTTGAGATCTCCGACACTCG 302

RESULT 9
LOCUS AA754458 247 bp mRNA EST 20-JAN-1998
DEFINITION 97SN1784 Rice Immature Seed Lambda ZAPII cDNA Library Oryza sativa
ACCESSION AA754458
NID 92801164
KEYWORDS EST.
SOURCE Rice.
ORGANISM Oryza sativa
REFERENCE 1 (bases 1 to 247)

```

AUTHORS Nahm, B.H., Kim, J.K., Cheong, J.J., Kim, S.I., Hahn, T.R., Moon, E.P., Kim, W.T., Kim, W.Y., Yang, M.S., Park, R.D., Sohn, U.I., Kang, K.Y., Lee, M.C. and Eun, M.Y.

JOURNAL Large-scale Sequencing Analysis of ESTs from Rice Immature Seed Unpublished (1998)

COMMENT Contact: Eun M.Y.
Department of Cytogenetics
National Inst. of Agri. Sci. and Tech, RDA
Suwon, Kyunggi-do, Korea
Tel: 82 331 290 0301
Fax: 82 331 290 0307
Email: myeun@n20.asti.re.kr
Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji University, Yongin, Korea. 449-728 bnhahm@bioserver.myongji.ac.kr
Seq primer: M13 Reverse Primer.

FEATURES Location/Qualifiers
1..247
/organism="Oryza sativa"
/cultivar="Milyang23"
/note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2: XhoI; Directional cDNA library inserted into lambda ZAPII vector at 5' end with EcoRI and 3' end with Xho I site."
/db_xref="taxon:4530"
/clone="97SN1784"
/clone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"
/tissue_type="Immature Seed"
/dev_stage="5 days after pollination"
/lab_host="E. coli SOLR"

BASE COUNT 7 a 16 c 21 g 34 t 169 others
ORIGIN

Query Match 3.2%; Score 55; DB 12; Length 247;
Best Local Similarity 14.9%; Pred. No. 3.18e-52;
Matches 34; Conservative 106; Mismatches 85; Indels 3; Gaps 3;

Db 1 HWDCTMTVWRCGCCBAMWN-KHFMHTBWBWCVRVCTTNNKNGKHGRITTTWDCSDN 59
Cp 374 CTTCGGAAGATCCCGGACCGGACATCTCGTCAAAACCCCTCGGTAGCTTGAGAT 315
Db 60 AHCRTVBYARYSGYGTBYISNVNDTGTGTGKTTVNVHSGWNNRCSNVVYVMB 119
Cp 314 CTC-CGACACTCGAAGTC-CTTAGGATAGTAGTACCGGTGAAGACCGATCCAGATGC 257
Db 120 TAYCYBHYBDRANVDDTRCTNDRGVCNYTASDNGTSATKRVYDKVTDSDCGGCGWCKR 179
Cp 256 CGCCACGCGCTGCCGGGTGTCGCAATGAAGGCGTGGTGCAGCTCTCTGCGCGCCAGA 197
180 VTGSSBVBRCGVNVVMTSMWTDKSKMBSMDXRRSRVHYGRWB 227
Cp 196 CTGAGAGCGGACCTCGCGCGGGGTGCTCCGCGCCACCGAG 149

RESULT 10
LOCUS T42569 352 bp mRNA EST 07-JAN-1998
DEFINITION 5832 Lambda-PRL2 Arabidopsis thaliana cDNA clone 114CI17T, mRNA sequence.
ACCESSION T42569
NID 92757838
KEYWORDS EST
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryotes; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; seed plants; Magnoliophyta;
eudicotyledons; Rosidae; Capparales; Brassicaceae; Arabidopsis.
1 (bases 1 to 352)
Newman, T., deBruijn, F.J., Green, P., Keegstra, K., Kendre, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomas, M., Retzel, E. and Somerville, C.
Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)
95148729

REFERENCE 1
AUTHORS Newman, T., deBruijn, F.J., Green, P., Keegstra, K., Kendre, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomas, M., Retzel, E. and Somerville, C.
TITLE Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones
JOURNAL Plant Physiol. 106, 1241-1255 (1994)
MEDLINE 95148729

AUTHORS Nahm, B.H., Kim, J.K., Cheong, J.J., Kim, S.I., Hahn, T.R., Moon, E.P., Kim, W.T., Kim, W.Y., Yang, M.S., Park, R.D., Sohn, U.I., Kang, K.Y., Lee, M.C. and Eun, M.Y.

JOURNAL Large-scale Sequencing Analysis of ESTs from Rice Immature Seed Unpublished (1998)

COMMENT Contact: Eun M.Y.
Department of Cytogenetics
National Inst. of Agri. Sci. and Tech, RDA
Suwon, Kyunggi-do, Korea
Tel: 82 331 290 0301
Fax: 82 331 290 0307
Email: myeun@n20.asti.re.kr
Submitted by Baek Hie Nahm, Dept of Biological Science, Myongji University, Yongin, Korea. 449-728 bnhahm@bioserver.myongji.ac.kr
Seq primer: M13 Reverse Primer.

FEATURES Location/Qualifiers
1..247
/organism="Oryza sativa"
/cultivar="Milyang23"
/note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2: XhoI; Directional cDNA library inserted into lambda ZAPII vector at 5' end with EcoRI and 3' end with Xho I site."
/db_xref="taxon:4530"
/clone="97SN1784"
/clone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"
/tissue_type="Immature Seed"
/dev_stage="5 days after pollination"
/lab_host="E. coli SOLR"

BASE COUNT 7 a 16 c 21 g 34 t 169 others
ORIGIN

Query Match 3.2%; Score 55; DB 12; Length 247;
Best Local Similarity 14.9%; Pred. No. 3.18e-52;
Matches 34; Conservative 106; Mismatches 85; Indels 3; Gaps 3;

Db 1 HWDCTMTVWRCGCCBAMWN-KHFMHTBWBWCVRVCTTNNKNGKHGRITTTWDCSDN 59
Cp 374 CTTCGGAAGATCCCGGACCGGACATCTCGTCAAAACCCCTCGGTAGCTTGAGAT 315
Db 60 AHCRTVBYARYSGYGTBYISNVNDTGTGTGKTTVNVHSGWNNRCSNVVYVMB 119
Cp 314 CTC-CGACACTCGAAGTC-CTTAGGATAGTAGTACCGGTGAAGACCGATCCAGATGC 257
Db 120 TAYCYBHYBDRANVDDTRCTNDRGVCNYTASDNGTSATKRVYDKVTDSDCGGCGWCKR 179
Cp 256 CGCCACGCGCTGCCGGGTGTCGCAATGAAGGCGTGGTGCAGCTCTCTGCGCGCCAGA 197
180 VTGSSBVBRCGVNVVMTSMWTDKSKMBSMDXRRSRVHYGRWB 227
Cp 196 CTGAGAGCGGACCTCGCGCGGGGTGCTCCGCGCCACCGAG 149

RESULT 10
LOCUS T42569 352 bp mRNA EST 07-JAN-1998
DEFINITION 5832 Lambda-PRL2 Arabidopsis thaliana cDNA clone 114CI17T, mRNA sequence.
ACCESSION T42569
NID 92757838
KEYWORDS EST
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryotes; Viridiplantae; Charophyta/Embryophyta group;
Embryophyta; Tracheophyta; seed plants; Magnoliophyta;
eudicotyledons; Rosidae; Capparales; Brassicaceae; Arabidopsis.
1 (bases 1 to 352)
Newman, T., deBruijn, F.J., Green, P., Keegstra, K., Kendre, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomas, M., Retzel, E. and Somerville, C.
Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)
95148729

COMMENT

On Jan 8, 1998 this sequence version replaced gi:933288.

Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313t@clmsu.edu
Seq primer: P7.

FEATURES

Location/Qualifiers

1..352

/organism="Arabidopsis thaliana"

/strain="var columbia"

/note="Vector: lambda Zip-Lox; Site_1: Sal; Site_2: Not; Lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is BRL's lambda Zip-Lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dt primed cDNA."

/db_xref="taxon:3702"

/clone="114CI17T"

/clone_lib="Lambda-PRL2"

BASE COUNT 76 a 83 c 75 g 104 t 14 others

ORIGIN

Query Match 3.0%; Score 51; DB 11; Length 352;

Best Local Similarity 65.4%; Pred. No. 3.01e-45;

Matches 106; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

Db 4 ATTCAAGGCAAGTCTACACGCTCCNATTGGATTAAACATCCCGAGCGACACG 63

QY 126 ATCCAGGCAAGTCTACACGCTCTCTCGGTGGCGGAGCAGCCCGCGGAGGTC 185

Db 64 GTGATTCTAAATCTCGTTGGTCAAGACGTCACCGATGCTTTCATCGCATTTATCCCGGA 123

QY 186 CGCTCTCAGTCTGGCGGCGGAGGAGTCACCGACGCTTCATTCGTCGACCCCGGCG 245

Db 124 ACCGTTGGCACCATCTCGACCATCTTTTCACCGGTTACAC 165

QY 246 ACGGCGTGGCGCATCTGTCGTCGCTCTTCACCGGCTACTAC 287

RESULT

LOCUS T15209 597 bp mRNA EST 28-JUL-1995

DEFINITION crs698 Ricinus communis cDNA clone pcrs698 similar to cytochrome

b5.

ACCESSION T15209

NID 9688663

KEYWORDS EST.

SOURCE

castor bean clone-pcrs698 library-lambdaZAPST strain-Baker 296 vector-lambdaZAPII primer-T3 Rsite1-EcoRI Rsite2-XhoI Poly(A)+ RNA was purified from developing stage III to stage V (Greenwood & Bewley, Can. J. Bot. 60:1751-1760, 1982) endosperm plus embryo of immature castor fruits. cDNA was synthesized and cloned into lambdaZAPII according to the instructions of the manufacturer (Stratagene); synthesis was primed from the poly(A) tail, and cloned directionally into XhoI (3') and EcoRI (5') sites. In few cases, sequence data indicated that this directionality was reversed. Partial cDNA clones predominate.

ORGANISM

Ricinus communis

Eucaryotes; Embryophyta; Magnoliophyta; Magnoliopsida; Rosidae;

Euphorbiales; Euphorbiaceae; Ricinus.

1 (bases 1 to 597)

AUTHORS van de Loo, F.J., Turner, S. and Somerville, C.

TITLE Expressed sequence tags from developing castor seeds

JOURNAL

Plant Physiol. 108, 1141-1150 (1995)

COMMENT

Db 1612 WTKWT-TWYMSMRWTMTTMTTWTWTWTWTWTWTWTWT 1650
 QY 1563 ATTTCCTATCCCATGCTATTTTGTGATTATGTTCTATT 1602

RESULT 14
 LOCUS A1107749 448 bp mRNA EST 26-AUG-1998
 DEFINITION GH05586.5prime GH Drosophila melanogaster head pot2 Drosophila
 melanogaster cDNA clone GH05586 5prime, mRNA sequence.

ACCESSION A1107749
 NID 93475402
 KEYWORDS EST.
 SOURCE fruit fly.

ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;

REFERENCE 1 (bases 1 to 448)

AUTHORS Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C.,
 Brokstein,P., Lewis,S. and Rubin,G.M.
 BDGP/HMI Drosophila EST Project
 UNpublished (1997)

FILE

JOURNAL

COMMENT

Contact: Harvey, D.
 G. M. Rubin-Molecular and Cell Biology
 University of California Berkeley
 539 LSA, Berkeley, CA 94720-3200, USA
 Fax: 510 643 9947
 Email: http://fruitfly.berkeley.edu/EST, est@fruitfly.berkeley.edu
 Plate: 55 row: H column: 2
 High quality sequence stop: 366.

Location/Qualifiers

1..448
 /organism="Drosophila melanogaster"
 /note="Organ: head; Vector: pot2; Site_1: EcoRI; Site_2:
 XhoI; Sized fractionated cDNAs were directly ligated into
 pot2. Plasmid cDNA library."
 /db_xref="taxon:7227"
 /clone="GH05586"
 /clone_lib="GH Drosophila melanogaster head pot2"
 /sex="male and female"
 /dev_stage="adult"
 /lab_host="DH5 - alpha"
 /lab a 120 c 126 g 91 t

BASE COUNT 111 a 120 c 126 g 91 t
 ORIGIN

Query Match 2.0%; Score 34; DB 17; Length 448;
 Best Local Similarity 62.5%; Pred. No. 1.05e-17;
 Matches 85; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Db 52 CACAACAAGCCAGGATCTGTGGTGTGTCATCGACAACAAGGTCTACGATGTGACCAAG 111
 QY 96 CACAACAAGTCGGCGATCTGTGGATCTCCATCCAGGCAAGGTCTACGACTGCTCTCGG 155

Db 112 TTCCGCTCTGAGCATCCCGTGGCGAGGAATCCCTGGTGGATGAGCGCGGTCCGATGCC 171
 QY 156 TGGCGCGGAGACACCCCGCGGCGAGGTCCCGCTCCTCAGTCTGGCGCGGAGGACGTC 215

Db 172 ACCAAGGCTTCAATG 187
 QY 216 ACCGACGCTTCATG 231

RESULT 15
 LOCUS A1107703 613 bp mRNA EST 26-AUG-1998
 DEFINITION GH05526.5prime GH Drosophila melanogaster head pot2 Drosophila
 melanogaster cDNA clone GH05526 5prime, mRNA sequence.

ACCESSION A1107703
 NID 93475356
 KEYWORDS EST.
 SOURCE fruit fly.

ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;
 Drosophilidae; Drosophila.

1 (bases 1 to 613)

AUTHORS Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C.,
 Brokstein,P., Lewis,S. and Rubin,G.M.
 BDGP/HMI Drosophila EST Project
 UNpublished (1997)

TITLE

JOURNAL

COMMENT

Contact: Harvey, D.
 G. M. Rubin-Molecular and Cell Biology
 University of California Berkeley
 539 LSA, Berkeley, CA 94720-3200, USA
 Fax: 510 643 9947
 Email: http://fruitfly.berkeley.edu/EST, est@fruitfly.berkeley.edu

Plate: 55 row: C column: 2

High quality sequence stop: 609.

Location/Qualifiers

1..613

/organism="Drosophila melanogaster"

/note="Organ: head; Vector: pot2; Site_1: EcoRI; Site_2:
 XhoI; Sized fractionated cDNAs were directly ligated into
 pot2. Plasmid cDNA library."
 /db_xref="taxon:7227"

/clone="GH05526"

/clone_lib="GH Drosophila melanogaster head pot2"

/sex="male and female"

/dev_stage="adult"

/lab_host="DH5 - alpha"

BASE COUNT 173 a 154 c 150 g 136 t
 ORIGIN

Query Match 2.0%; Score 34; DB 17; Length 613;

Best Local Similarity 62.5%; Pred. No. 1.05e-17;

Matches 85; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Db 52 CACAACAAGCCAGGATCTGTGGTGTGTCATCGACAACAAGGTCTACGATGTGACCAAG 111
 QY 96 CACAACAAGTCGGCGATCTGTGGATCTCCATCCAGGCAAGGTCTACGACTGCTCTCGG 155

Db 112 TTCCGCTCTGAGCATCCCGTGGCGAGGAATCCCTGGTGGATGAGCGCGGTCCGATGCC 171
 QY 156 TGGCGCGGAGACACCCCGCGGCGAGGTCCCGCTCCTCAGTCTGGCGCGGAGGACGTC 215

Db 172 ACCAAGGCTTCAATG 187
 QY 216 ACCGACGCTTCATG 231

Search completed: Thu Feb 18 17:40:14 1999
 Job time : 4650 secs.

MAEPLH (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsearch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run: Thu Feb 18 17:50:01 1999; MasPar time 122.19 Seconds
1204.694 Million cell updates/sec

Tabular output not generated.

Title: >US-08-934-254-26

Description: (1-1702) from US08934254.seq

Perfect Score: 1702
N.A. Sequence: 1 CCCCAAAATTTTCATTGT.....TTTTGGTAAAAA 1702
Comp: GGGGTTTTAAAGTAACAA.....AAAAACATTTTTTTTT

Scoring table: TABLE default

Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 165359 seqs, 43243793 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-issued
1:5A_COMB 2:5B_COMB 3:5C_COMB 4:PCT9_COMB 5:backfiles1

Statistics: Mean 9.108; Variance 5.151; scale 1.768

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	172	10.1	1685	1	US-08-366-Sequence 4, Applicatio	5.25e-102
2	172	10.1	1685	3	US-08-789-Sequence 4, Applicatio	5.25e-102
3	91	5.3	7218	2	US-08-232-Sequence 14, Applicati	5.86e-44
4	46	2.7	9653	3	US-08-388-Sequence 22, Applicati	7.85e-14
5	40	2.4	9653	3	US-08-388-Sequence 22, Applicati	3.99e-10
6	37	2.2	7218	2	US-08-232-Sequence 14, Applicati	2.52e-08
7	36	2.1	2151	1	US-08-238-Sequence 5, Applicatio	9.86e-08
8	34	2.0	2151	1	US-08-238-Sequence 5, Applicatio	1.45e-06
9	27	1.6	74	4	PCT-US95-1 Sequence 94, Applicati	1.16e-02
10	28	1.6	81	4	PCT-US95-1 Sequence 92, Applicati	3.36e-03
11	27	1.5	242	2	US-08-273-Sequence 1, Applicatio	1.16e-02
12	26	1.5	74	4	PCT-US95-1 Sequence 94, Applicati	3.89e-02
13	25	1.5	74	4	PCT-US95-1 Sequence 100, Applicati	1.28e-01
14	26	1.5	75	4	PCT-US95-1 Sequence 99, Applicati	3.89e-02
15	26	1.5	81	4	PCT-US95-1 Sequence 92, Applicati	3.89e-02
16	25	1.5	81	4	PCT-US95-1 Sequence 98, Applicati	1.28e-01
17	25	1.5	82	4	PCT-US95-1 Sequence 97, Applicati	1.28e-01
18	26	1.5	82	4	PCT-US95-1 Sequence 97, Applicati	3.89e-02
19	25	1.5	92	3	US-08-353-Sequence 16, Applicati	1.28e-01
20	25	1.5	242	2	US-08-273-Sequence 1, Applicatio	1.28e-01

21	25	1.5	2016	1	US-08-294-Sequence 1, Applicatio	1.28e-01
22	25	1.5	2016	4	PCT-US95-0 Sequence 1, Applicatio	1.28e-01
c 23	26	1.5	5852	1	US-07-867-Sequence 2, Applicatio	3.89e-02
c 24	23	1.4	66	1	US-08-471-Sequence 144, Applicat	1.31e+00
25	23	1.4	69	1	US-08-471-Sequence 142, Applicat	1.31e+00
c 26	23	1.4	69	1	US-08-471-Sequence 142, Applicat	1.31e+00
27	23	1.4	74	4	PCT-US95-1 Sequence 100, Applicat	1.31e+00
28	24	1.4	75	4	PCT-US95-1 Sequence 99, Applicati	4.14e-01
29	23	1.4	76	3	US-08-478-Sequence 95, Applicati	1.31e+00
30	23	1.4	76	1	US-08-120-Sequence 98, Applicati	1.31e+00
31	23	1.4	81	4	PCT-US95-1 Sequence 30, Applicati	1.31e+00
32	23	1.4	90	3	US-08-442-Sequence 6, Applicatio	4.14e-01
33	24	1.4	165	4	PCT-US93-0 Sequence 1, Applicatio	1.31e+00
34	24	1.4	1662	3	US-08-651-Sequence 1, Applicatio	1.31e+00
35	23	1.4	2830	4	PCT-US93-0 Sequence 1, Applicatio	1.31e+00
36	23	1.4	2830	1	US-07-882-Sequence 10, Applicatio	1.31e+00
c 37	23	1.4	5181	3	US-08-257-Sequence 25, Applicati	4.01e+00
c 38	22	1.3	84	4	PCT-US95-1 Sequence 120, Applicat	4.01e+00
c 39	22	1.3	91	2	US-08-133-Sequence 4, Applicatio	4.01e+00
c 40	22	1.3	84	4	PCT-US95-0 Sequence 1, Applicatio	4.01e+00
c 41	22	1.3	768	3	US-08-567-Sequence 10, Applicati	4.01e+00
c 42	22	1.3	2088	1	US-08-073-Sequence 3, Applicatio	4.01e+00
43	22	1.3	2504	2	US-08-483-Sequence 1, Applicatio	4.01e+00
44	22	1.3	2626	1	US-08-156-Sequence 1, Applicatio	4.01e+00
45	22	1.3	49377	2	US-08-764-Sequence 1, Applicatio	4.01e+00

ALIGNMENTS

RESULT 1
ID US-08-366-779-4 STANDARD; DNA; UNC; 1685 BP.
AC xxxxxx

Sequence 4, Application US/08366779

Sequence 4, Application US/08366779

Patent No. 5614393

GENERAL INFORMATION:

APPLICANT: Thomas, Terry L.

APPLICANT: Reddy, Avutu S.

APPLICANT: Nuccio, Michael

APPLICANT: Freytsinet, Georges L.

APPLICANT: Nunberg, Andrew N.

TITLE OF INVENTION: PRODUCTION OF GAMMA LINOLENIC ACID BY A

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESS:

ADDRESSEE: Scully, Scott, Murphy & Presser

STREET: 400 Garden City Plaza

CITY: Garden City

STATE: New York

COUNTRY: United States

ZIP: 11530

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/366,779

FILING DATE: 30-DEC-1994

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Presser, Leopold

REGISTRATION NUMBER: 19,827

REFERENCE/DOCKET NUMBER: 8383ZYXW

TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 742-4343

TELEFAX: (516) 742-4366

TELEX: 230 901 SANS UR

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 1685 base pairs

TYPE: nucleic acid

CC	STRANDEDNESS: both	
CC	TOPOLOGY: linear	
CC	MOLECULE TYPE: DNA (genomic)	
SQ	SEQUENCE 1685 Bp; 431 A; 277 C; 357 G; 620 T; 0 OTHER.	
Query Match	10.1%; Score 172; DB 1; Length 1695;	
Best Local Similarity	57.7%; Pred. No. 5.25e-102;	
Matches	697; Conservative 0; Mismatches 507; Indels 3; Gaps	
Db	42 CAATGGCTGCTCAAAATCAAGAAATACATTACCTCAGATGAACCTCAAGACACACGACATAAAC	101
QY		
Db	46 CAATGAGGCGGAAGCTAAGATGATATACCGCGGAGGAGCCTCCGCGCCCAACAACAGT	105
Db	102 CCGGAGATCTATGGATCTCGATTCAAGGGAAGCCCTATGATCTTCGGATTGGGTGAAG	161
QY		
Db	106 CCGGCGATCTCTGGATCTCCATCCAGGGCAAGGTCTACGACTGCTCTCGGTGGGCGGCG	165
Db	162 ACCATCCAGGTGCGACGTTTCCTTTGAAGAGTCTTGTCTGGTCAAGAGGTAACTGATGAT	221
QY		
Db	166 AGCACCCCGCGGAGGTCCGCTCCTCAGTCTTGCGCGCCAGGACGTCACCGACGCT	225
Db	222 TTGTTGCATTCACCTCGCTCTACATGAAAGAACTCTGATPAAGTTTTTCACTGGGTATT	281
QY		
Db	226 TCATTGCGTACACCGCGGACGGGTGGCGCATCTGGATCCGCTCTTCAACCGCTACT	285
Db	282 A-T-CCTAAAGATTACTCTGTTCTGAGGTTTCTAAAGATTATAGGAGCTTGTGTTG	338
QY		
Db	286 ACTACTCAGGACTTCGAAGTGTGGAGATCTCCAGGACTACCGGAGGCTTTTGAACG	345
Db	339 AGTTTCTAAATGGGTTTGTATGACAAAAAGGTCATATTATGTTGCACTTTTGCT	398
QY		
Db	346 AGATGTCGGTCCGGGATCTTCGAGAAGAGGCCACCATCATGTGGAGTTCGTGCG	405
Db	399 TTATACAAATGCTTGTGCTATGAGTGTTTATGAGGGTTTGTGTTGTGAGGGTGTTTGG	458
QY		
Db	406 GCGTTGCGGTATGATGGCGGAATCGTCTACGGCGTGTGGCGTGGAGTCGGTCGGAG	465
Db	459 TACATTGTTTCTCGGTTGTTGATGGGGTTTCTTTGATTCAGACTGGTGGATTGGAC	518
QY		
Db	466 TTCATGCTCTGCGCGCACCTGCTGGGCTTGTCTGTGATCCAGCCCGTATGTGGGCC	525
Db	519 ATGATGCTGGGCATTATATGTTAGTGTCTGATTCAAGGCTTAATAAGTTTATGGGTATTT	578
QY		
Db	526 ATGACTCCGGCATTTACAGAGTGATGCCAACCCGTGGATACACAGAAATCAGCAACTCA	585
Db	579 TTGCTGCAATTTGCTTTCAGGAATAAGTATTGTTGGTGGAAATGGAACCATATATGCAC	638
QY		
Db	586 TAGCAGGCAACATCCTTAACCGGAATCAGCATCGCGTGTGGAGTGGACCCACACAGCCC	645
Db	639 ATCATAATGCCGTGTAATAGCCTTGAATATGACCCTGATTTACAATATATACATTCCTTG	698
QY		
Db	646 ACCACCTGCCCTGCAACAGCCTCGACTACGACCCCGAGCTCCAGCACATCCCGCTATTTCG	705
Db	699 TTGTGCTTCCAGTTTTTTGGTTTCACTCACCTCTCATTTCTATGAGAAAGGTTGACTT	758
QY		
Db	706 CGCTCTCCACCGACTTTTCAACTCCATCACCTCGGCTCTTATGGCGGAGTCTCTGAAAT	765
Db	759 TTGACTCTTTATCAAGATCTTTGTAAGTATCAACATTGGACATTTTACCTCTATTATGT	818
QY		
Db	766 TCGACGAGTGGCAGGTTCTTAGTCAGCTACCAGCACTGGACCTACTACCCGGTCAATGA	825
Db	819 GTGCTGCTAGGCTCAATATGATGTACAATCTCTCATATGTGTGTGACCAGAGAAATG	878
QY		
Db	826 TCTTGGCGGAGTCAACCTTTTCATCCAGACCTTTTTTATTGCTCTCCTCACAGGCGCAGC	885
Db	879 TGTCTATCGACTCAGGAACCTTTGGGAGCCGTAGTGTCTCGATTTGTTGACCGGTTCG	938
QY		
Db	886 TCCCTGACCGGCTCTAACTTTAATGGGTATCGCGGTTTTCTGGACGTGGTTCCTCGCTCT	945
Db	939 TTGTTTCTTTGTTGCTTAATTTGGGTGAAAGAAATATATGTTTGTATTGCAAGTTTATCAG	998
QY		
Db	946 TGTATFTTGTCTCCGAACTGGCTGAACGGTTGGGTTCTGCTCTCATACAGTTTCGGG	1005

Db	999	TGACTGGAAATGCAACAAGTTTCAGTTCTCCTTTGAACCACTTCCTTTCAAGTGTTCATGTTG	10505
Qy	1006	TCAGGGCGATCCAGCACGTCAGTTCCAGCTCAACCACTTCCTCCGGCGACACATACGTG	1065
Db	1059	GAAGCCCTAAAGGGAATTAATTTGGTTTGAGAAACAACGATGGGACACTTGACATTTCTT	1118
Qy	1066	GCCCCCAAGGGCGACAACTGGTTTCGAGAGACAGCAACGAAAGGACGATTCGATATCACGT	1125
Db	1119	GTCCTCTTGATGGATTGGTTTCATGTGGATTCGAATTCCAAATTGAGCATCATTTCT	1178
Qy	1126	GCCACCGTGGATGGACTGGTTCTTTGGTGGCTCAGTTCCAGTTGGAGCACCACCTGT	1185
Db	1179	TTCCCAAGATGCCATAGATGCAACCTTAGGAAAAATCTCGCCCTACGTGATCGAGTTATGCA	1238
Qy	1186	TCCCTAGCTGCCGCGTGGCGAGCTTAGAAGATTCGCGCCCTTGCTCGGACTTGTGTGA	1245
Db	1239	AGAAACA 1245	
Qy	1246	AGAACA 1252	
RESULT	2		
ID	US-08-789-936-4	STANDARD; DNA; UNC; 1685 BP.	
AC	xxxxxx		
DT	Sequence 4, Application US/08789936		
DE	Sequence 4, Application US/08789936		
CC	Patent No. 5789220		
CC	GENERAL INFORMATION:		
CC	APPLICANT: Thomas, Terry L.		
CC	APPLICANT: Reddy, Avutu S.		
CC	APPLICANT: Nuccio, Michael		
CC	APPLICANT: Freyssinet, Georges L.		
CC	APPLICANT: Nunberg, Andrew N.		
CC	TITLE OF INVENTION: PRODUCTION OF GAMMA LINOLENIC ACID BY A		
CC	TITLE OF INVENTION: DELTA 6-DESATURASE		
CC	NUMBER OF SEQUENCES: 25		
CC	CORRESPONDENCE ADDRESS:		
CC	ADDRESSEE: Scully, Scott, Murphy & Presser		
CC	STREET: 400 Garden City Plaza		
CC	CITY: Garden City		
CC	STATE: New York		
CC	COUNTRY: United States		
CC	ZIP: 11530		
CC	COMPUTER READABLE FORM:		
CC	MEDIUM TYPE: Floppy disk		
CC	COMPUTER: IBM PC compatible		
CC	OPERATING SYSTEM: PC-DOS/MS-DOS		
CC	SOFTWARE: Patentin Release #1.0, Version #1.25		
CC	CURRENT APPLICATION DATA:		
CC	APPLICATION NUMBER: US/08/789,936		
CC	FILING DATE: 28-JAN-1997		
CC	CLASSIFICATION: 435		
CC	PRIOR APPLICATION DATA:		
CC	APPLICATION NUMBER: 08/366,779		
CC	FILING DATE: 30-DEC-1994		
CC	ATTORNEY/AGENT INFORMATION:		
CC	NAME: Presser, Leopold		
CC	REGISTRATION NUMBER: 19,827		
CC	REFERENCE/DOCKET NUMBER: 8383ZYXW		
CC	TELECOMMUNICATION INFORMATION:		
CC	TELEPHONE: (516) 742-4343		
CC	TELEFAX: (516) 742-4366		
CC	TELEX: 230 901 SANS UR		
CC	INFORMATION FOR SEQ ID NO: 4:		
CC	SEQUENCE CHARACTERISTICS:		
CC	LENGTH: 1685 base pairs		
CC	TYPE: nucleic acid		
CC	STRANDEDNESS: both		
CC	TOPOLOGY: linear		
CC	MOLECULE TYPE: DNA (genomic)		
CC	SEQUENCE 1685 BP; 431 A; 277 C; 357 G; 620 T; 0 OTHER.		

Db	776	GGRKKGURHUHVHSGVRSTSTCTASDYYTTSYWGWRGWDYGGGYTNYNGKRGRVTM	835
		: : : : : : : : : : : : : : : : : : : :	
Cp	798	GGTAGTCCTAGGAACCGTGCCACTTCTCGAATTTCAAGACTCGGCCATAAGAAGCCG	739
Db	836	ADTSSNSRSVTAADTVAVYCVRGRRYSDDGDWGGTTVTVVSHUVDKMTSSSASVGD	895
		: : : : : : : : : : : : : : : : : : : : :	
Cp	738	AGGTGATGGAGTAGGAGTCGGGTGGACGCACAATACGGGGATGTCTCGAGGTCGG	679
Db	896	RVITCRSSITHGNNTYYWKAKAYRNVNRSGVSRSGSGSTDYTTSSDAHYC	950
		: : : : : : : : : : : : : : : : : : : : :	
Cp	678	GGT-CGTGTCFAGAGCTGTTGACAGCGAGAGTGGTGGCGCTTGTGGCTCACATTCC	625

RESULT 6
ID US-08-232-463-14 STANDARD; DNA; UNC; 7218 BP.
AC xxxxxx

Sequence 14, Application US/08232463
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFELINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
City: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463

CC FILING DATE: CC
CC CLASSIFICATION: 435 CC
CC PRIOR APPLICATION DATA: CC
CC APPLICATION NUMBER: US/07/935,313 CC
CC FILING DATE: CC
CC APPLICATION NUMBER: EP 91 114 300.6 CC
CC FILING DATE: 26-AUG-1991 CC
CC ATTORNEY/AGENT INFORMATION: CC
CC NAME: BENT, Stephen A. CC
CC REGISTRATION NUMBER: 29,768 CC
CC REFERENCE/DOCKET NUMBER: 30472/114 IMMU CC
CC TELECOMMUNICATION INFORMATION: CC
CC TELEPHONE: (703)836-9300 CC
CC TELEFAX: (703)683-4109 CC
CC TELEX: 899149 CC
CC INFORMATION FOR SEQ ID NO: 14: CC
CC SEQUENCE CHARACTERISTICS: CC
CC LENGTH: 7218 base pairs CC
CC TYPE: nucleic acid CC
CC STRANDEDNESS: single CC
CC TOPOLOGY: linear CC
CC IMMEDIATE SOURCE: CC
CC CLONE: pTZ9pt-F15 CC
CC SEQUENCE 7218 BP: 1944 A: 1491 C: 1486 G: 1929 T: 368 OTHER. CC
CC SQ

```

Query Match      2.2%   Score 37;   DB 2;   Length 7218;
Best Local Similarity 0.8%   Pred. No. 2.52e-08;
Matches 3;   Conservative 197;   Mismatches 163;   Indels 0;   Gaps 0;

Db 1064  GATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1123
||||:||||:   ::::   :   :   :   :   :   :   :   :   :   :   :   :   :
Cp 1436  GATCCCTTCAACCCACCAACCAACCAACCAATCAACCATGGGTGTGTATTAAGCTT 1377
||||:||||:   ::::   :   :   :   :   :   :   :   :   :   :   :   :

```

[illegible]

RESULT 7
ID US-08-238-163-5 STANDARD; DNA; UNC; 215 BP.
AC xxxxxx

Sequence 5, Application US/08238163
Sequence 5, Application US/08238163
Patent No. 5569830
GENERAL INFORMATION:

CC APPLICANT: BENNETT, Alan
CC APPLICANT: LABAVITCH, John M.

CC APPLICANT: POWELL, Ann
CC APPLICANT: STOTZ, Henrik

CC TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL

CC	TITLE OF INVENTION
CC	NUMBER OF SEQUENCES

CC NUMBER OF SEQUENCES: 2
CC CORRESPONDENCE ADDRESS:

CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend

CC STREET: Steuart Street

CITY: San Francisco

CC STATE: California
CC COUNTRY: US

COUNTRY: US
ZIP: 94105-1493

CC
CC
CC

CC MEDIUM TYPE: Floppy

CC COMPUTER: IBM PC COMP

CC OPERATING SYSTEM: PC
CC SOFTWARE: Patent In R
CC

CC	SOFTWARE: FALCONLIN K
CC	CURRENT APPLICATION DAT

CC APPLICATION NUMBER:

CC FILING DATE: 03-MAY-

CC CLASSIFICATION: 800
CC ATTORNEY/AGENT INFORMATION

CC ATTORNEY/AGENT INFORMATION
CC NAME: Bastian, Kevin

CC NEED: DUSTMAN, REVIN
CC REGISTRATION NUMBER:
CC

CC	REFERENCE/DOCKET NUMB	NUMB
CC	REFERENCE/DOCKET NUMB	NUMB

CC TELECOMMUNICATION INFORMATION

TELEPHONE: (415) 543-5433
TELEFAX: (415) 543-5433

TELEFAX: (415) 343-5

CC INFORMATION FOR SEY ID NO
CC SEQUENCE CHARACTERISTIC
CC

CC LENGTH: 215 base pai

CC TYPE: nucleic acid

CC STRANDEDNESS: single
CC TOPOLOGY: unknown

CC TOPOLOGY: unknown
CC MOLECULE TYPE: protein

[illegible]

NAME/KEY: misc_featu

CC LOCATION: 1..215
CC OTHER INFORMATION: /standard_name= "Deduced amino acid
CC OTHER INFORMATION: sequence of PCIP from bean."
SQ SEQUENCE 215 BP; 15 A; 8 C; 25 G; 26 T; 141 OTHER.

Query Match 2.1%; Score 36; DB 1; Length 215;
Best Local Similarity 16.7%; Pred. No. 9.86e-08;
Matches 31; Conservative 69; Mismatches 85; Indels 1; Gaps 1;

Db 10 SVVSRATSCNDKAKKGGTSSWTTCCNTGWCVDTTTTRVNNDSGHKNKYSSANYNG 69
Cp 955 CAAGATACGAAGAGGGAACCGCTCCAGAAACCGCATACCATTAAGTTTAGAGCG 896
Db 70 GNNVGAATHTYHTNVSADSKTVDTSYNASGTSSSNGGT-DGNRSGADSYGSSKTAAT 128
Cp 895 CGGTCAGGAGCGTCCGCTGGTGAGGAGCAATAAAAGGTCTGGATGAAGAGTTGACT 836
Db 129 SRNRTGTANNVDSRNMGDASVSKNTKKHAKNSADGKVGSKNNNGDRNNRYGTGKSN 188
Cp 835 CGCCGGAAGATCATGACCGGTAGTAGTCCAGTCTGGTAGCTAGTACTAGGACCGTGC 776

189 VSNNGC 194
:: ||
775 ACTTCG 770

RESULT 8
ID US-08-238-163-5 STANDARD; DNA; UNC; 215 BP.
AC xxxxxx
DT

Sequence 5, Application US/08238163
Sequence 5, Application US/08238163
CC Patent No. 5569830
CC GENERAL INFORMATION:
CC APPLICANT: BENNETT, Alan
CC APPLICANT: LABAVITCH, John M.
CC APPLICANT: POWELL, Ann
CC APPLICANT: STOTZ, Henrik
CC TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL
CC TITLE OF INVENTION: POLYCALACTURONASES AND THEIR USE TO CONTROL FUNGAL DISEASE
CC NUMBER OF SEQUENCES: 24
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend Kourie and Crew
CC STREET: Stewart Street Tower, One Market Plaza
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: US
CC ZIP: 94105-1493
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/238,163
CC FILING DATE: 03-MAY-1994
CC CLASSIFICATION: 800
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Bastian, Kevin L.
CC REGISTRATION NUMBER: 34,774
CC REFERENCE/DOCKET NUMBER: 2307E-540
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 543-9600
CC TELEFAX: (415) 543-5043
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 215 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: protein
CC FEATURE:
CC NAME/KEY: misc_feature

CC LOCATION: 1..215
CC OTHER INFORMATION: /standard_name= "Deduced amino acid
CC OTHER INFORMATION: sequence of PCIP from bean."
SQ SEQUENCE 215 BP; 15 A; 8 C; 25 G; 26 T; 141 OTHER.

Query Match 2.0%; Score 34; DB 1; Length 215;
Best Local Similarity 13.0%; Pred. No. 1.45e-06;
Matches 19; Conservative 60; Mismatches 67; Indels 0; Gaps 0;

Db 45 DTDTTYRVNDSGHKNKYSSANYNGNNVGAATHTYHTNVSADSKTVDTSYNASGTS 104
QY 1145 GTTCTTTGGTGGCTGCAGTTCAGTTGGAGCACCACTTGTTCCTAGCTGCCGCGTGG 1204
Db 105 SSSGTTGDSGADSYGSSKTAATSRNRTGTANNVDSRNMGDASVSGDNTKKHAKNS 164
QY 1205 GCACCTTAGGAAGATTGCGCCCTTGGCTCGGAGCTTGTGTAAAGACGACGGATGCCGTA 1264
Db 165 ADGKVGSKNNNGDRNNRYGTGKSNVS 190
QY 1265 TAGGAGCTTCGGTTTGGGACGAGC 1290

RESULT 9
ID PCT-US95-11934-94 STANDARD; DNA; UNC; 74 BP.
AC xxxxxx
DT

Sequence 94, Application PC/TUS9511934
Sequence 94, Application PC/TUS9511934
CC GENERAL INFORMATION:
CC APPLICANT: Cytogen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From
CC TITLE OF INVENTION: Peptide Libraries
CC NUMBER OF SEQUENCES: 103
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Misrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 94:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 74 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC SEQUENCE 74 BP; 3 A; 4 C; 3 G; 1 T; 63 OTHER.

Query Match 1.6%; Score 27; DB 4; Length 74;
Best Local Similarity 10.1%; Pred. No. 1.16e-02;
Matches 7; Conservative 20; Mismatches 42; Indels 0; Gaps 0;

Db 3 GAGNN 62
Cp 686 GAGGTCGGGTCGTAGTCGAGGCTGTTGCAGGCGAGGTGTTGGGCTTGGGTCACCTT 627

Db 63 BNNBNACGC 71
: : : :
Cc 626 CCACCACGC 618

RESULT 10
ID PCT-US95-11934-92 STANDARD; DNA; UNC; 81 BP.
AC xxxxxx
DT
DE Sequence 92, Application PC/TUS9511934
Cc Sequence 92, Application PC/TUS9511934
Cc GENERAL INFORMATION:
Cc APPLICANT: Cytogen Corporation
Cc TITLE OF INVENTION: Antigen Binding Peptides (Abitides) From
Cc TITLE OF INVENTION: Peptide Libraries
Cc NUMBER OF SEQUENCES: 103
Cc CORRESPONDENCE ADDRESS:
Cc ADDRESSEE: Pennie & Edmonds
Cc STREET: 1155 Avenue of the Americas
Cc CITY: New York
Cc STATE: New York
Cc COUNTRY: USA
Cc ZIP: 10036
Cc COMPUTER READABLE FORM:
Cc MEDIUM TYPE: Floppy disk
Cc COMPUTER: IBM PC compatible
Cc OPERATING SYSTEM: PC-DOS/MS-DOS
Cc SOFTWARE: Patent In Release #1.0, Version #1.30
Cc CURRENT APPLICATION DATA:
Cc APPLICATION NUMBER: PCT/US95/11934
Cc FILING DATE: 20-SEP-1995
Cc CLASSIFICATION:
Cc ATTORNEY/AGENT INFORMATION:
Cc NAME: Misrock, S. Leslie
Cc REGISTRATION NUMBER: 18,872
Cc REFERENCE/DOCKET NUMBER: 1101-196-228
Cc TELECOMMUNICATION INFORMATION:
Cc TELEPHONE: (212) 790-9090
Cc TELEFAX: (212) 869-9741/8864
Cc TELEX: 66141 PENNIE
Cc INFORMATION FOR SEQ ID NO: 92:
Cc SEQUENCE CHARACTERISTICS:
Cc LENGTH: 81 base pairs
Cc TYPE: nucleic acid
Cc STRANDEDNESS: single
Cc TOPOLOGY: linear
Cc MOLECULE TYPE: DNA (genomic)
Cc SEQUENCE 81 BP; 3 A; 5 C; 6 G; 4 T; 63 OTHER.

Query Match 1.6%; Score 28; DB 4; Length 81;
Best Local Similarity 12.5%; Pred. No. 3.36e-03;
Matches 9; Conservative 20; Mismatches 43; Indels 0; Gaps 0;

Db 7 CTCGAGNNBN 66
||| :
Cc 689 CTGAGTCGGGTCTGCTAGTCAGGCTGTTGCAGCGAGTGTGTGGGTCCA 630
:
Cc 67 BNNBNBNACGC 78
: : : : :
Cc 629 CTTCCACACGC 618

RESULT 11
ID US-08-273-846-1 STANDARD; DNA; UNC; 290 BP.
AC xxxxxx
DT
DE Sequence 1, Application US/08273846
Cc Sequence 1, Application US/08273846
Cc Patent No. 5641863
Cc GENERAL INFORMATION:
Cc APPLICANT: SCHREIBER, ALAN D.
Cc APPLICANT: PARK, JONG-GU

Db 63 BNNBNACGC 71
: : : :
Cc 626 CCACCACGC 618

RESULT 10
ID PCT-US95-11934-92 STANDARD; DNA; UNC; 81 BP.
AC xxxxxx
DT
DE Sequence 92, Application PC/TUS9511934
Cc Sequence 92, Application PC/TUS9511934
Cc GENERAL INFORMATION:
Cc APPLICANT: Cytogen Corporation
Cc TITLE OF INVENTION: Antigen Binding Peptides (Abitides) From
Cc TITLE OF INVENTION: Peptide Libraries
Cc NUMBER OF SEQUENCES: 103
Cc CORRESPONDENCE ADDRESS:
Cc ADDRESSEE: Pennie & Edmonds
Cc STREET: 1155 Avenue of the Americas
Cc CITY: New York
Cc STATE: New York
Cc COUNTRY: USA
Cc ZIP: 10036
Cc COMPUTER READABLE FORM:
Cc MEDIUM TYPE: Floppy disk
Cc COMPUTER: IBM PC compatible
Cc OPERATING SYSTEM: PC-DOS/MS-DOS
Cc SOFTWARE: Patent In Release #1.0, Version #1.30
Cc CURRENT APPLICATION DATA:
Cc APPLICATION NUMBER: PCT/US95/11934
Cc FILING DATE: 20-SEP-1995
Cc CLASSIFICATION:
Cc ATTORNEY/AGENT INFORMATION:
Cc NAME: Misrock, S. Leslie
Cc REGISTRATION NUMBER: 18,872
Cc REFERENCE/DOCKET NUMBER: 1101-196-228
Cc TELECOMMUNICATION INFORMATION:
Cc TELEPHONE: (212) 790-9090
Cc TELEFAX: (212) 869-9741/8864
Cc TELEX: 66141 PENNIE
Cc INFORMATION FOR SEQ ID NO: 92:
Cc SEQUENCE CHARACTERISTICS:
Cc LENGTH: 81 base pairs
Cc TYPE: nucleic acid
Cc STRANDEDNESS: single
Cc TOPOLOGY: linear
Cc MOLECULE TYPE: DNA (genomic)
Cc SEQUENCE 81 BP; 3 A; 5 C; 6 G; 4 T; 63 OTHER.

Query Match 1.6%; Score 28; DB 4; Length 81;
Best Local Similarity 12.5%; Pred. No. 3.36e-03;
Matches 9; Conservative 20; Mismatches 43; Indels 0; Gaps 0;

Db 7 CTCGAGNNBN 66
||| :
Cc 689 CTGAGTCGGGTCTGCTAGTCAGGCTGTTGCAGCGAGTGTGTGGGTCCA 630
:
Cc 67 BNNBNBNACGC 78
: : : : :
Cc 629 CTTCCACACGC 618

RESULT 11
ID US-08-273-846-1 STANDARD; DNA; UNC; 290 BP.
AC xxxxxx
DT
DE Sequence 1, Application US/08273846
Cc Sequence 1, Application US/08273846
Cc Patent No. 5641863
Cc GENERAL INFORMATION:
Cc APPLICANT: SCHREIBER, ALAN D.
Cc APPLICANT: PARK, JONG-GU

Db 63 BNNBNACGC 71
: : : :
Cc 626 CCACCACGC 618

RESULT 10
ID PCT-US95-11934-92 STANDARD; DNA; UNC; 81 BP.
AC xxxxxx
DT
DE Sequence 92, Application PC/TUS9511934
Cc Sequence 92, Application PC/TUS9511934
Cc GENERAL INFORMATION:
Cc APPLICANT: Cytogen Corporation
Cc TITLE OF INVENTION: Antigen Binding Peptides (Abitides) From
Cc TITLE OF INVENTION: Peptide Libraries
Cc NUMBER OF SEQUENCES: 103
Cc CORRESPONDENCE ADDRESS:
Cc ADDRESSEE: Pennie & Edmonds
Cc STREET: 1155 Avenue of the Americas
Cc CITY: New York
Cc STATE: New York
Cc COUNTRY: USA
Cc ZIP: 10036
Cc COMPUTER READABLE FORM:
Cc MEDIUM TYPE: Floppy disk
Cc COMPUTER: IBM PC compatible
Cc OPERATING SYSTEM: PC-DOS/MS-DOS
Cc SOFTWARE: Patent In Release #1.0, Version #1.30
Cc CURRENT APPLICATION DATA:
Cc APPLICATION NUMBER: PCT/US95/11934
Cc FILING DATE: 20-SEP-1995
Cc CLASSIFICATION:
Cc ATTORNEY/AGENT INFORMATION:
Cc NAME: Misrock, S. Leslie
Cc REGISTRATION NUMBER: 18,872
Cc REFERENCE/DOCKET NUMBER: 1101-196-228
Cc TELECOMMUNICATION INFORMATION:
Cc TELEPHONE: (212) 790-9090
Cc TELEFAX: (212) 869-9741/8864
Cc TELEX: 66141 PENNIE
Cc INFORMATION FOR SEQ ID NO: 92:
Cc SEQUENCE CHARACTERISTICS:
Cc LENGTH: 81 base pairs
Cc TYPE: nucleic acid
Cc STRANDEDNESS: single
Cc TOPOLOGY: linear
Cc MOLECULE TYPE: DNA (genomic)
Cc SEQUENCE 81 BP; 3 A; 5 C; 6 G; 4 T; 63 OTHER.

Query Match 1.6%; Score 28; DB 4; Length 81;
Best Local Similarity 12.5%; Pred. No. 3.36e-03;
Matches 9; Conservative 20; Mismatches 43; Indels 0; Gaps 0;

Db 7 CTCGAGNNBN 66
||| :
Cc 689 CTGAGTCGGGTCTGCTAGTCAGGCTGTTGCAGCGAGTGTGTGGGTCCA 630
:
Cc 67 BNNBNBNACGC 78
: : : : :
Cc 629 CTTCCACACGC 618

RESULT 11
ID US-08-273-846-1 STANDARD; DNA; UNC; 290 BP.
AC xxxxxx
DT
DE Sequence 1, Application US/08273846
Cc Sequence 1, Application US/08273846
Cc Patent No. 5641863
Cc GENERAL INFORMATION:
Cc APPLICANT: SCHREIBER, ALAN D.
Cc APPLICANT: PARK, JONG-GU

Db 63 BNNBNACGC 71
: : : :
Cc 626 CCACCACGC 618

RESULT 10
ID PCT-US95-11934-92 STANDARD; DNA; UNC; 81 BP.
AC xxxxxx
DT
DE Sequence 92, Application PC/TUS9511934
Cc Sequence 92, Application PC/TUS9511934
Cc GENERAL INFORMATION:
Cc APPLICANT: Cytogen Corporation
Cc TITLE OF INVENTION: Antigen Binding Peptides (Abitides) From
Cc TITLE OF INVENTION: Peptide Libraries
Cc NUMBER OF SEQUENCES: 103
Cc CORRESPONDENCE ADDRESS:
Cc ADDRESSEE: Pennie & Edmonds
Cc STREET: 1155 Avenue of the Americas
Cc CITY: New York
Cc STATE: New York
Cc COUNTRY: USA
Cc ZIP: 10036
Cc COMPUTER READABLE FORM:
Cc MEDIUM TYPE: Floppy disk
Cc COMPUTER: IBM PC compatible
Cc OPERATING SYSTEM: PC-DOS/MS-DOS
Cc SOFTWARE: Patent In Release #1.0, Version #1.30
Cc CURRENT APPLICATION DATA:
Cc APPLICATION NUMBER: PCT/US95/11934
Cc FILING DATE: 20-SEP-1995
Cc CLASSIFICATION:
Cc ATTORNEY/AGENT INFORMATION:
Cc NAME: Misrock, S. Leslie
Cc REGISTRATION NUMBER: 18,872
Cc REFERENCE/DOCKET NUMBER: 1101-196-228
Cc TELECOMMUNICATION INFORMATION:
Cc TELEPHONE: (212) 790-9090
Cc TELEFAX: (212) 869-9741/8864
Cc TELEX: 66141 PENNIE
Cc INFORMATION FOR SEQ ID NO: 92:
Cc SEQUENCE CHARACTERISTICS:
Cc LENGTH: 81 base pairs
Cc TYPE: nucleic acid
Cc STRANDEDNESS: single
Cc TOPOLOGY: linear
Cc MOLECULE TYPE: DNA (genomic)
Cc SEQUENCE 81 BP; 3 A; 5 C; 6 G; 4 T; 63 OTHER.

Query Match 1.6%; Score 28; DB 4; Length 81;
Best Local Similarity 12.5%; Pred. No. 3.36e-03;
Matches 9; Conservative 20; Mismatches 43; Indels 0; Gaps 0;

Db 7 CTCGAGNNBN 66
||| :
Cc 689 CTGAGTCGGGTCTGCTAGTCAGGCTGTTGCAGCGAGTGTGTGGGTCCA 630
: : : : ~ :
Cc 67 BNNBNBNACGC 78
: : : : :
Cc 629 CTTCCACACGC 618

RESULT 11
ID US-08-273-846-1 STANDARD; DNA; UNC; 290 BP.
AC xxxxxx
DT
DE Sequence 1, Application US/08273846
Cc Sequence 1, Application US/08273846
Cc Patent No. 5641863
Cc GENERAL INFORMATION:
Cc APPLICANT: SCHREIBER, ALAN D.
Cc APPLICANT: PARK, JONG-GU

Db 63 BNNBNACGC 71
: : : :
Cc 626 CCACCACGC 618

RESULT 10
ID PCT-US95-11934-92 STANDARD; DNA; UNC; 81 BP.
AC xxxxxx
DT
DE Sequence 92, Application PC/TUS9511934
Cc Sequence 92, Application PC/TUS9511934
Cc GENERAL INFORMATION:
Cc APPLICANT: Cytogen Corporation
Cc TITLE OF INVENTION: Antigen Binding Peptides (Abitides) From
Cc TITLE OF INVENTION: Peptide Libraries
Cc NUMBER OF SEQUENCES: 103
Cc CORRESPONDENCE ADDRESS:
Cc ADDRESSEE: Pennie & Edmonds
Cc STREET: 1155 Avenue of the Americas

CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mirock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELEX: 56141 PENNIE
CC INFORMATION FOR SEQ ID NO: 94:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 74 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
SQ SEQUENCE 74 BP: 3 A; 4 C; 3 G; 1 T; 63 OTHER.

```
Query Match 1.5%; Score 26; DB 4; Length 74;
Best Local Similarity 8.8%; Pred. No. 3.89e-02;
Matches 6; Conservative 20; Mismatches 42; Indels 0; Gaps 0;
```

[illegible]

D _b	65	NBNACGCC	72
		:	
Q _Y	217	CCGACGCC	224

RESULT 13
ID PCT-US95-11934-100 STANDARD; DNA; UNC; 74 BP.
AC xxxxxx

DT	Sequence	100, Application	PC/TUS9511934
DE	Sequence 100,	Application	PC/TUS9511934
CC	Sequence 100,	Application	PC/TUS9511934

CC GENERAL INFORMATION:
CC APPLICANT: Cytogen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From
CC TITLE OF INVENTION: Peptide Libraries
CC NUMBER OF SEQUENCES: 103
CC CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036

DATE: 10030
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/11934
FILING DATE: 20-SEP-1995

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-196-228
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864

CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 100:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 74 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC SEQUENCE 74 BP; 6 A,6 C; 1 G; 1 T; 60 OTHER.

Query Match 1.5%; Score 25; DB 4; Length 74;
Best Local Similarity 9.0%; Pred. No. 1.28e-01;
Matches 6; Conservative 20; Mismatches 41; Indels 0; Caps 0;

Db 1 CTAGAVNNVNVNVNNTVVNNVNVTNVNVNVNVNVNVNVNVNVNVNVNVNV 60
||| |: : : : : : : : : : : : : :
Cp 1267 CTATACGCCATCCCGTGCTTCTTACACAAGTCCCAGGCCAAGGGGCGCAATCATTCCTAAGC 1208

Db	61	NNVNAC	67
		:	
Cp	1207	TGCCAC	1201

RESULT 14
ID PCT-US95-11934-99 STANDARD; DNA; UNC; 75 BP.
AC xxxxxx

AC	xxxxxxx
DT	
DE	Sequence 99, Application PC/TUS9511934
CC	Sequence 99, Application PC/TUS9511934
CC	GENERAL INFORMATION:

CC	GENERAL INFORMATION:
CC	APPLICANT: CytoGen Corporation
CC	TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From
CC	TITLE OF INVENTION: Peptide Libraries
CC	NUMBER OF SEQUENCES: 103
CC	CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036

CC ZIP: 10036
CC
CC COMPUTER READABLE FORM:
CC
CC MEDIUM TYPE: Floppy disk
CC
CC COMPUTER: IBM PC compatible
CC
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC
CC CURRENT APPLICATION DATA:
CC
CC APPLICATION NUMBER: PCT/US95/11934
CC
CC FILING DATE: 20-SEP-1995
CC

DATE: 20 SEP 1993
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-196-228
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE

CC	TELEX: 66141 PENNLE	99:
CC	INFORMATION FOR SEQ ID NO:	
CC	SEQUENCE CHARACTERISTICS:	
CC	LENGTH: 75 base pairs	
CC	TYPE: nucleic acid	
CC	STRANDEDNESS: single	
CC	TOPOLOGY: linear	
CC	MOLECULE TYPE: DNA (genomic)	
CC	SEQUENCE 75 BP: 1 A: 1 C: 7 G: 5	

Query Match 1.5%; Score 26; DB 4; Length 75;
Best Local Similarity 9.0%; Pred. No. 3.89e-02;
Matches 6; Conservative 20; Mismatches 41; Indels

[illegible]

Search completed: Thu Feb 18 17:54:01 1999
Job time : 240 secs.